

Vurdering som profesjonsfaglig kompetanse og som verktøy for læring

*- Et innblikk i vurderingskulturen i to barneskoler i Edmonton, hovedstaden i
Alberta, Canada.*

TRUDE SLEMMEN



Hovedoppgave i profesjonsstudiet i pedagogikk
Undervisning og ledelse

Våren 2006

UNIVERSITETET I OSLO
Det utdanningsvitenskaplige fakultet
Pedagogisk forskningsinstitutt

SAMMENDRAG

TITTEL:

Vurdering som profesjonsfaglig kompetanse og som verktøy for læring. – Et innblikk i vurderingskulturen i to barneskoler i Edmonton, Alberta.

AV:

SLEMMEN, Trude

SEMESTER:

Vår 2006

EKSAMEN:

Hovedoppgave i profesjonsstudiet i pedagogikk

2. avdeling, undervisning og ledelse

Cand. ed.

1. PROBLEMOMRÅDE

Målet med denne oppgaven er å se hvordan vurderingskulturen i Alberta har påvirket et utvalg læreres bruk av vurdering som profesjonsfaglig kompetanse og som verktøy for læring i to barneskoler i hovedstaden Edmonton. Alberta har hatt provinsielle prøver for elever på 6., 9. og 12. trinn siden 1982 og dette har satt sitt preg på vurderingskulturen i provinsen fordi resultatvurdering har hatt et stort fokus. På bakgrunn av dette har jeg følgende problemstillinger: 1) *Hvordan preger Albertas vurderingskultur læreres bruk av vurdering som profesjonsfaglig kompetanse og som verktøy for læring i to barneskoler i Edmonton?* 2) *Hvordan kan erfaringene fra Edmonton berike den norske vurderingsdebatten?*

2. TEORIGRUNNLAG

Begrepet profesjonsfaglig kompetanse knytter jeg til Dale (1999) som vektlegger fire kunnskapselementer som utgjør lærerens profesjonsgrunnlag. Disse er kunnskap, fagdidaktikk, kollegialitet og pedagogiske kriterier for kvalitet. I denne oppgaven vil jeg knytte begrepet ”profesjonsfaglig” til disse fire elementene og. I analysen vil jeg bruke disse fire kunnskapselementene som analyseredskap. Begrepet ”profesjonsfaglig” vil jeg også knytte til forskningsprosjektet ”Vurdering som profesjonsfaglig kompetanse” (Haugstveit, Sjølie og Øygarden 2006). Dette prosjektet har som formål å undersøke et utvalg norske læreres praktiske vurderingsferdighet og gi innsikt i hvordan vurderingskompetansen blir forstått av et utvalg norske lærere på 5., 6. og 7. trinn. Ved hjelp av en empirisk undersøkelse, overfører jeg ”vurdering som profesjonsfaglig kompetanse” til Edmonton. I to barneskoler som utfører to ulike vurderingspraksiser vil jeg undersøke et utvalg læreres forståelse av vurderingsbegrepet.

For å forklare hvordan vurdering kan brukes som verktøy for læring i klasserommet tar jeg utgangspunkt i Black og Wiliam (1998). Black og Wiliam ser på vurdering som: ”all those activities undertaken by teachers, and by their students, which provide information to be used as feedback to modify the teaching and learning activities in which they are engaged” (Black og Wiliam 1998a: 2). På bakgrunn av denne definisjonen tar jeg utgangspunkt i to ulike måter å vurdere på: ”vurdering FOR læring” og ”vurdering AV læring”. ”Vurdering FOR læring” vil si å innsamle informasjon om elevens læringsutbytte eller prosess med formålet om å forbedre læring. ”Vurdering AV læring” vil si å vurdere et læringsutbytte på

et gitt tidspunkt (jf AAC¹ 2005). Begrepene ”skolekode” (jf Arfwerdson) og rammefaktor vil også være sentrale i oppgaven.

For å redegjøre for vurderingsdebatten i Norge tar jeg utgangspunkt i nasjonale og internasjonale undersøkelser som sier noe om den norske vurderingskulturen. Klettess (2003) analyse av Reform 97, Elevinspektørene 2005, prosjektet ”Vurdering som profesjonsfaglig kompetanse” (2006) og en pilotundersøkelse jeg gjorde i forkant av mitt studie i Edmonton, er eksempler på dette. Videre tar jeg for meg historiske og kulturelle trekk ved utdanningen i Norge som kan være med på å forklare den norske vurderingskulturen.

3. METODE

Datamateriale og empiri i denne oppgaven er anskaffet ved hjelp av kvalitative intervju av lærere, observasjoner og innsamling av skriftlige dokumenter i fem barneskoler i Edmonton (skole A, B, C, D og E). Jeg tar utgangspunkt i skole A og B, og vektlegger derfor det empiriske materialet fra disse to skolene i min analyse. I tillegg til studiene i Edmonton vil jeg bruke funn fra pilotundersøkelsen i to barneskoler i Oslo (skole F og skole G). Jeg tar også i bruk komparativ metode ved å sammenligne vurderingskulturen i skole A og B. I flere sammenhenger gjør jeg også en komparativ sammenligning av vurderingskulturen i Norge og Alberta.

4. HOVEDKONKLUSJONER

I analysen i del 6 tar jeg for meg problemstilling nr. 1: *Hvordan preger Albertas vurderingskultur læreres bruk av vurdering som profesjonsfaglig kompetanse og som verktøy for læring i to barneskoler i Edmonton?* Ved å ta utgangspunkt i intervjuer fra disse to skolene, gjør jeg en analyse av hvordan vurderingskulturen i Alberta preger lærernes vurderingspraksis i de to skolene. Skole A er et eksempel på en skole som overfokuserer bruken av ”vurdering AV læring”, noe som kan føre til *ineffektiv læring* (jf Black og Wiliam 1998a). I denne skolen samlet lærerne inn informasjon om elevens kompetanse hovedsaklig ved hjelp av summative prøver. Vurderingsfokuset og deres retningslinjer for ”vurdering AV læring” begrenset lærerens handlingsrom til å bruke ulike varierte vurderingsmetoder underveis i læringsprosessen. Skole B fokuserer på ”vurdering FOR læring” i stor grad, ved siden av ”vurdering AV læring”. Med andre ord samlet lærerne inn informasjon om elevens

¹ Alberta Assessment Consortium

kompetanse både ved hjelp av summative og formative vurderingsstrategier. Videre knyttet lærerne disse to vurderingsformene tett sammen i skole B ved at de brukte tydelige kriterier som elevene var bevisste. Dette kan i følge Black og Wiliam (1998b) være noe som kan føre til *effektiv læring* ved at eleven er inkludert i større grad. Et slikt vurderingsfokus ga lærerne i skole B et større handlingsrom til å bruke ulike vurderingsmetoder. Sammenligningen av skole A og B tyder på at ”skolekoden” styrer lærernes handlingsrom i større grad enn læreplanen og nasjonale prøver, ved at lærerne tolker læreplanen og bruken av de provinsielle prøvene ut fra skolens filosofi, ledelse og fastlagte rammer.

I del 7 knytter jeg erfaringene fra mitt empiriske studie i Edmonton til den norske vurderingsdebatten for å besvare problemstilling nr. 2: *Hvordan kan erfaringene fra Edmonton berike den norske vurderingsdebatten?* For å tydeliggjøre vurderingsdebatten som pågår i Norge i dag, tar jeg utgangspunkt Utdanningsdirektoratets beskrivelser av formål og dilemmaer med en eventuell innføring av nasjonale vurderingskriterier. Erfaringene fra skole A og B tyder på at ”vurdering AV læring” og tydelige kriterier kan føre til både negative og positive konsekvenser.

Forord

I Norge ble nasjonale prøver gjennomført første gang våren 2004, men har av ulike årsaker blitt utsatt til skoleåret 2007/ 2008. Det har vært mye diskusjoner blant pedagoger og i medier om disse prøvene vil fungere eller ikke, og mange har vært skeptiske til en slik vurderingsform. Da denne vurderingsdebatten pågikk våren 2005, befant jeg meg i Edmonton, hovedstaden i Albertaprovinsen i Canada, der jeg tilbrakte et semester av profesjonsstudiet i pedagogikk ved Universitetet i Alberta. Som student i Alberta opplevde jeg en sterk vurderingskultur med stort fokus på kriteriesetting, vurdering og tilbakemeldinger fra professorer og medstudenter. På Universitetet hendte det at flere av de Canadiske studentene stilte spørsmål som ”hva må jeg ha med i oppgaven for å få en A -?”, ” hva kreves for å få en A+?” Den samme vurderingskulturen opplevde jeg også i de barne- og ungdomsskolene jeg besøkte i Edmonton. Elever helt ned til fem års alderen fikk karakterer og lærerne jobbet mye med å sette vurderingskriterier for og sammen med elevene. Dette var en uvant vurderingskultur for meg som i alle år har vært preget av den norske vurderingskulturen. Ved Universitetet i Oslo har kritisk refleksjon og selvstendig didaktisk tenkning vært en viktig kompetanse blant professorer og studenter. Som norsk skoleelev ble jeg ikke kjent med begrepet karakterer før jeg begynte på ungdomsskolen. Jeg har heller aldri fulgt en bestemt ”oppskrift” for å få den karakteren jeg har ønsket meg. I starten var jeg kritisk til den vurderingskulturen jeg opplevde i Alberta fordi det var ukjent for meg. Jeg tenkte at et slikt vurderingsfokus kunne føre til en ”teaching to the test” mentalitet og overfokusering på resultater. Etter hvert som tiden gikk ble jeg usikker på om konsekvensene bare var negative. Som student opplevde jeg at fokuset på kriteriesetting og vurdering fungerte som et godt virkemiddel for meg i min læringsprosess. Jeg visste til enhver tid hva som ble forventet av meg og jeg fikk tydelige og klare tilbakemeldinger fra professorer og medstudenter.

Ved siden av å være en del av en annerledes vurderingskultur i Edmonton enn jeg var vant til i Norge, var det også noe annet som vekket min interesse i Alberta. I PISA 2003 var Alberta den provinsen som gjorde det best nasjonalt og lå tett på Finlands gode resultater på flere områder (jf vedlegg 1). Alberta har vært preget av en sterk tradisjon for vurdering av læringsresultater og jeg tenkte at dette kunne være en årsak. Jeg tenkte at en eventuell ”teaching to the test” undervisning ville føre til gode resultater, men at det kunne gå på

bekostning av lærernes profesjonelle kompetanse. For å få en bedre forståelse av vurderingskulturen i Alberta dro jeg tilbake til Edmonton fire uker i mars 2006.

Takk til "Alberta Education", "Edmonton Public School Board", "Alberta Assessment Consortium", "Alberta Teacher Association" og "University of Alberta", som hjalp meg med å bli kjent med vurderingskulturen i Alberta. Takk til elever, lærere og rektorer på de fem barneskolene jeg besøkte i Edmonton og de to barneskolene jeg besøkte i Oslo. En stor takk til min hovedveileder Britt Ulstrup Engelsen, som har vært med meg under hele arbeidet. Takk til Hans Petter Wille ved Høyskolen i Østfold, Yngvild Nilsen ved Utdanningssetaten i Oslo, Sverre Tveit, Elisabeth Thomsen fra "Edmonton Public School Board" og Helene Fulland for gjennomlesning og innspill. Spesielt takk til min medstudent Karen Klyve Sunde som også har bidratt til gjennomlesning og vært en god diskusjonspartner i Canada. En stor takk til avdeling for pedagogisk utvikling og kvalitet i Utdanningssetaten i Oslo som har latt meg få delta i deres utviklingsarbeid i vurderingsfeltet i Osloskolen parallelt som jeg har skrevet min oppgave. Astrid Søggen, Elin Reite, Yngvild Nilsen, Marthe Amundsen og øvrige medarbeidere i avdeling for pedagogisk utvikling og kvalitet, har vært gode støttespillere i min skriveprosess. Til sist en stor takk til min gode kjæreste Frank, min søster og mine venner som har støttet meg i tiden jeg har jobbet med denne oppgaven.

Oslo, 1. september 2006

Trude Slemmen

INNHOLD

1. INNLEDNING	11
1.1 TEORIGRUNNLAG	13
1.1.1 Problemstilling.....	14
1.2 AKTUALISERING OG MOTIV	15
2 HVA ER VURDERING?	18
2.1 “ASSESSMENT” VERSUS “EVALUATION”	18
2.2 FORMATIV OG SUMMATIV VURDERING	20
2.3 VURDERING FOR LÆRING I KLASSEROMMET	24
2.4 VURDERINGSKRITERIER	27
2.5 VURDERING SOM PROFESJONSFAGLIG KOMPETANSE.....	30
3 METODEKAPITTEL	33
3.1 VITENSKAPSTEORETISK RAMME	33
3.2 KVANTITATIV OG KVALITATIV METODE.....	34
3.3 PLANLEGGING AV TEMA, UTVALG OG METODER	34
3.4 KOMPARATIV METODE	37
3.5 DET KVALITATIVE FORSKNINGINTERVJU	39
3.6 OBSERVASJON	40
3.7 VALIDITET, RELIABILITET OG GENERALISERBARHET	41
4 VURDERINGSKULTUREN I ALBERTA	43
4.1 UTDANNINGSPOLITISKE TREKK	43
4.2 BEHOV FOR “ACCOUNTABILITY” I ALBERTA	45
5 PRESENTASJON AV FUNN	47
5.1 SKOLE A: VURDERING AV LÆRING I FOKUS.	47

5.1.1	<i>Observasjoner tatt fra en klasstime</i>	48
5.1.2	<i>Observasjoner tatt fra et foreldre – lærer- intervju.</i>	49
5.2	SKOLE B: VURDERING FOR LÆRING I FOKUS	50
5.2.1	<i>Observasjoner tatt i en klasstime</i>	51
5.2.2	<i>Observasjoner fra en "student led conference"</i>	51
5.3	ØVRIGE OBSERVASJONER I SKOLE C, D OG E	51
6	ANALYSE	54
6.1	VURDERING SOM PROFESJONSFAGLIG KOMPETANSE	54
6.1.1	<i>Hva slags kunnskap har lærerne om vurdering?</i>	54
6.1.2	<i>Hvordan spesifiseres kriterier, målene og innhold for vurderingen?</i>	57
6.1.3	<i>Finnes det noen felles retningslinjer for vurdering i lærerkollegiet?</i>	59
6.1.4	<i>Hvordan utvikles kriterier for vurdering som skal fremme kvalitet i skolen?</i>	62
6.2	VURDERING SOM VERKTØY FOR LÆRING I KLASSEROMMET	63
6.3	HVA SIER INTERVJUENE OM "SKOLEKODEN" PÅ SKOLE A OG SKOLE B?	65
6.4	OPPSUMMERING	67
7	HVORDAN KAN ERFARINGENE FRA EDMONTON BERIKE DEN NORSKE VURDERINGSDEBATTEN?	69
7.1	TREKK VED VURDERINGSKULTUREN I NORGE.....	69
7.1.1	<i>Behov for ansvarliggjøring også i Norge</i>	70
7.1.2	<i>Vurdering som verktøy for læring i norske klasserom</i>	71
7.2	HVA KAN VI LÆRE AV ERFARINGENE FRA SKOLE A OG B?	75
7.2.1	<i>Samsvar mellom summativ og formativ vurdering</i>	76
7.2.2	<i>Tydighet og kriteriesetting</i>	77
7.2.3	<i>Vurdering kan føre til konkurranse og utenatføring</i>	78

7.2.4	<i>Vurdering som en selvstendig fagdisiplin</i>	79
7.2.5	<i>Ledelsen bør delta i vurderingsarbeidet</i>	80
7.2.6	<i>Vurdering behøver ikke nødvendigvis bety mistillit til lærerprofesjonaliteten</i>	81
8	AVSLUTTENDE KOMMENTAR	83
	VEDLEGG 1: FUNN FRA PISA 2000 / 2003	86
	VEDLEGG 2: INTERVJUGUIDE	90
	VEDLEGG 3: MODELL	92
	REFERANSER	93

1. Innledning

Med sine 3.306 000 innbyggere² ligger Alberta vest i Canada. Hovedstaden Edmonton har 712.391 innbyggere³ og er Albertas nest største by etter Calagary. Alberta fikk sin provinsielle status i 1905 og er den rikeste provinsen i Canada på grunn av olje og naturgass produksjon. Jordbruk og kvegdrift har også en viktig rolle med hensyn til Albertas økonomi⁴. Klimaet og naturen i Alberta ligner det norske. De har fire årstider der vinteren er preget av kulde og mye snø. Fjellkjeden Rocky Mountains strekker seg over vestre delen av Alberta, noe som gjør at ski og friluftsliv er viktige aktiviteter for innbyggerne i provinsen. Til tross for flere likheter når det gjelder klima og økonomiske ressurser i Alberta og Norge, skiller Alberta seg fra Norge på flere områder. Deres ulike vurderingskultur er et av dem. For å få en forståelse av hvorfor vurderingskulturene har utviklet seg i ulik retning ser jeg det som viktig å plassere Alberta og Norge i en utdanningspolitisk kontekst. Med bakgrunn i dette, vil jeg se hvordan Albertas vurderingskultur preger vurderingspraksisen i to barneskoler Edmonton og se hvordan erfaringene fra disse skolene kan berike vurderingsdebatten i Norge.

Alberta og resten av Canada var rammet av en stor økonomisk og politisk krise på begynnelsen av 1970 tallet og skolen fikk skylden for dette. Med målet om å heve kunnskapsnivået ble nyliberale⁵ strategier som satte effektivitet og ansvarliggjøring i fokus innført. Denne utdanningspolitiske omveltningen har satt sitt preg på utdanningssystemet i Alberta, helt fram til i dag (Kachur 1999). I Norge har den politiske utviklingen vært veldig ulik utviklingen i Alberta. I følge Telhaug (1992 i Engelsen 2003) har den sosialdemokratiske modellen, til tross for justeringer som er kommet mot slutten av 1980 årene, i det store og det hele beholdt sitt særpreg i Norge. ”Ideen om skolen som et instrument for å fremme sosial rettferdighet har fremdeles en sterkere posisjon i Norge enn de fleste andre land vi sammenlikner oss med” (Telhaug 1992 i Engelsen 2003: 63). Med andre ord har vi ikke hatt det samme behovet for å øke kunnskapsnivået av økonomiske årsaker i Norge som i Alberta.

² Per år 2005

³ Per år 2005

⁴ <http://en.wikipedia.org/wiki/Alberta>

⁵ Jf. David Harveys definisjon på nyliberalisme s. 44

En forklaring på hvorfor det amerikanske samfunnet er mer følsomt for nyliberalisme enn det norske, kan i følge Telhaug være at de tilhører to ulike velferdsstater. For å forklare dette henviser han til Gosta Esping – Anderson som i en studie i 1990, ”Three Worlds of Welfare Capitalism”, redegjorde for tre ulike velferdsstater (i Telhaug 1994). Ifølge denne modellen er Norge en del av den sosialdemokratiske velferdsstaten som først og fremst finnes i de skandinaviske land. Denne velferdsstaten fremmer en likhet av høyeste standard, de offentlige tjenester holdes på et høyt nivå, og de sikrer samtidig at arbeiderklassen blir delaktig i de rettigheter som tilkommer de bedrestilte. Modellen inkluderer alle borgere i et universelt trygdesystem, og den konstruerer derfor en universell solidaritet til fordel for velferdsstaten (Telhaug 1994: 19). Canada derimot, er en del av den liberale velferdsstaten der markedstenkningen og den tradisjonelle liberale arbeidsmoralen står sterkt og det legges derfor vekt på at velferdstiltakene begrenses slik at det ikke oppstår misbruk (Telhaug 1994:19). Denne inndelingen kan være en årsak til at Canada og Norge har en ulik utdanningspolitisk fortid, noe som preger utdanningspolitikken i landene og dermed skolens innhold og praksisformer.

Et annet viktig moment som kan ha påvirket de to ulike vurderingskulturene i Alberta og Norge, er deres ulike læreplantradisjoner. Ifølge Gundem og Hopman (2002) kan disse to tradisjonene kalles for: ”the anglo – saxon tradition of curriculum studies” og “the North european tradition of didactic”. Disse modellene hviler på to ulike tradisjoner når det gjelder lærerutdanning og nasjonale – lokale forhold. Den første modellen som kan oversettes til den amerikanske læreplantradisjonen er metodestyrte i stor grad og blir sett på som en selvstendig yrkesfaglig disiplin. I denne tradisjonen lærer lærerstudenter metoder og teknikker for god undervisning og vurdering. Dette er i kontrast til den nordeuropeiske didaktikktradisjonen som verdsetter den autonome og reflekterte læreren. Læreren skal selv danne sin undervisningsteori og metoder på bakgrunn av refleksjon og egen fortolkning av læreplanen (Gundem og Hopman 2002). Ut fra denne inndelingen, kan man si at Norge tilhører nordeuropeiske didaktikktradisjonen, mens Alberta tilhører den amerikanske læreplantradisjonen. Alberta har en vurderingskultur som er metodestyrte i stor grad og som er preget av sentrale vurderingssystemer. Videre har vurderingsteori og metodikk vært en viktig del av lærerutdanningen. Norge derimot, er i en startfase når det kommer til å etablere et nasjonalt vurderingssystem. Videre har vurderingsmetodikk ikke hatt en tydelig rolle i den norske lærerutdanningen fram til nå (jf del 2.4).

1.1 Teorigrunnlag

Gjennomgående i hele oppgaven vil jeg ta utgangspunkt i det engelske ordet "assessment" som jeg oversetter til vurdering. I Norge har begrepene vurdering og evaluering ofte blitt brukt med samme betydning (jf Engelsen 2006, Lysne 1999 og Gundem 1998) og jeg vil i denne oppgaven skille disse begrepene fra hverandre for å tydeliggjøre vurderingsbegrepet (jf del 2.1). Jeg vil ta utgangspunkt i Black og Wiliams definisjon på vurdering der vurdering er: "all those activities undertaken by teachers, and by their students, which provide information to be used as feedback to modify the teaching and learning activities in which they are engaged" (Black og Wiliam 1998a: 2). Evaluering vil jeg forklare som en bedømmelse vedrørende kvalitet og verdi (jf Mervin (1968), Stake (1969) og Venezky (1974) i Lysne 1999:307).

Til Black og Wiliams definisjon av vurdering, knytter jeg begrepene "assessment for learning" og "assessment of learning". Begrepet "assessment for learning" som jeg oversetter til "vurdering FOR læring", vil si å innsamle informasjon om elevens læringsutbytte eller prosess med formålet om å forbedre læring. Begrepet "assessment of learning" oversetter jeg til "vurdering AV læring", som inkluderer vurdering av et læringsutbytte på et bestemt tidspunkt (AAC 2005). "Vurdering FOR læring" kan sees sammenheng med formativ vurdering som vektlegger vurderingen underveis i prosessen for å oppnå et best mulig sluttresultat. "Vurdering AV læring" kan sees i sammenheng med summativ vurdering som blir sett som en avsluttende vurdering av et sluttprodukt (jf Engelsen 2006: 121).

For å forklare begrepet profesjonsfaglig kompetanse, tar jeg utgangspunkt i Dale (1999) som vektlegger fire kunnskapselementer som utgjør lærerens profesjonsgrunnlag. Disse er kunnskap, fagdidaktikk, kollegialitet og pedagogiske kriterier for kvalitet. I min analyse av hvordan lærere bruker vurdering som profesjonsfaglig kompetanse i Edmonton blir det tatt utgangspunkt i fire spørsmål som bygger på disse elementene. Videre knyttes vurderingsbegrepet til prosjektet "Vurdering som profesjonsfaglig kompetanse" som ble gjort av Høyskolen i Hedemark i forbindelse med forskningsprogrammet, KUPP⁶ høsten 2001. Praksisnærhet, faglighet, tverrfaglighet og metodisk mangfold er nøkkelbegrep som knyttes til profesjonsbegrepet i KUPP prosjektet (Hoel 2005) og jeg vil også knytte disse

begrepene til ”profesjonsfaglig kompetanse”. På bakgrunn av dette prosjektet og Dales profesjonsbegrep, vil jeg se hvordan vurdering brukes som profesjonsfaglig kompetanse og som verktøy for læring i to barneskoler i Edmonton. For å forklare hvordan vurdering kan brukes som verktøy for læring i klasserommet, tar jeg utgangspunkt i Black og Wiliam (1998a). Black og Wiliams teori blir oppgavens hovedteori for å forklare viktigheten av å fokusere på vurderingen som skjer underveis i læringsprosessen. De vektlegger egenvurdering og gir eksempler på hvordan man kan se formative strategier i sammenheng med summative strategier for å oppnå læring. Begrepene ”skolekode” (jf Arfwerdson) og rammefaktorer vil også være sentrale i analysen.

1.1.1 Problemstilling

Med utgangspunkt i min empiriske undersøkelse i Albertas hovedstad Edmonton, har jeg følgende problemstillinger:

- 1) Hvordan preger Albertas vurderingskultur læreres bruk av vurdering som profesjonsfaglig kompetanse og som verktøy for læring i to barneskoler i Edmonton?*
- 2) Hvordan kan erfaringene fra Edmonton berike vurderingsdebatten i Norge?*

For å redegjøre for problemstilling nummer 1, anvendes følgende empiriske metoder: analysen baserer seg på kvalitative intervjuer med til sammen seks lærere på 5. og 6. trinn og observasjoner i to barneskoler i Edmonton (skole A og B). I tillegg har jeg gjort observasjoner i tre andre barneskoler i Edmonton (skole C, D og E). Jeg har også hatt samtaler med pedagogiske og politiske støttespillere som har gitt meg informasjon om vurderingens rolle og omfang i Alberta og Edmonton. Jeg har snakket med informanter fra ”Alberta Assessment Consortium” (AAC), lærerutdanningen i Alberta, ”Alberta Teacher’s Association” (ATA)⁷, ”Alberta Education” (utdanningsdepartementet), vurderingsavdelingen i ”Edmonton Public School Board” (EPSB) og rektorer på fem barneskoler og hvor en er en offentlig støttet ”charterskole”⁸. Jeg vil gjennomgående i oppgaven omtale ”Edmonton Public School Board” som EPSB. Internasjonale og nasjonale vurderingsrapporter og politiske dokumenter som sier noe om vurderingskulturen i Norge og Alberta, har også vært sentrale i oppgaven.

⁶ ”Kunnskapsutvikling i profesjonsutdanning og profesjonsutøving”

⁷ Lærerorganisasjonen i Alberta

⁸ Se definisjon i del 5 s. 47

Begrepet *kultur* vil jeg definere som: “a set of solutions devised by a group of people to meet specific problems posed by situations they face in common (....) This notion of culture as a living, historical product group problem solving allows an approach to cultural study that is applicable to any group, be it society, a neighbourhood, a family, a dance band, or an organization and its segments” (Van Maanen og Barley i Rubin og Rubin 1995:20). Jeg vil ta utgangspunkt i denne definisjonen for å forklare hvordan kultur er et historisk produkt som preger menneskers handlinger. En slik oppfattelse av kultur begrepet er i tråd med min vitenskapsteoretiske oppfatning som er sosialkonstruktivistisk der mennesker påvirker og påvirkes av den kultur de lever i (jf del 3.1). Albertas utdanningspolitiske utvikling og historie som siden tidlig på 1980 tallet har vært preget av nyliberalisme og effektivitet i stor grad, har preget vurderingskulturen i skolene i Alberta og Edmonton.

For å besvare problemstilling nummer 2: *Hvordan kan erfaringene fra Edmonton berike vurderingsdebatten i Norge?* tar jeg for meg noen interessante funn fra internasjonale og nasjonale undersøkelser som gir et innblikk i den norske vurderingskulturen. Disse undersøkelsene vil jeg se i sammenheng med seks intervjuer jeg har gjort i Oslo. Før jeg dro til Edmonton gjorde jeg en pilotstudie, der jeg intervjuet seks lærere i to barneskoler i Oslo (skole F og G). Disse lærerne ble intervjuet med utgangspunkt i den samme intervjuguiden jeg brukte i Edmonton. Svarene fra disse intervjuene analyseres ikke, men brukes som eksempler på hvordan et lite utvalg av norske lærere på 5. og 6. trinn i Oslo tenker rundt sin kunnskap om vurdering. Avslutningsvis vil jeg diskutere hvordan erfaringene fra Edmonton kan belyse den norske vurderingsdebatten blant annet ved å ta utgangspunkt i 6 punkter som jeg har utformet for å tydeliggjøre interessante erfaringer fra Edmonton.

1.2 Aktualisering og motiv

Det som vekket min interesse i Alberta var først og fremst deres vurderingskultur som er preget av resultatvurdering i stor grad, og deres usedvanlige gode resultater både i PISA undersøkelsen 2000 og 2003. I PISA 2000 gjorde elever i Alberta det best av alle OECD landene i vurdering av leseferdigheter og lå som nummer tre i matematikk (Se figur 3 og 4 i vedlegg 1). I PISA 2003 lå Alberta blant de beste OECD landene i matematikk (se figur 7 vedlegg 1). I PISA 2003 lå Alberta rett under Finland, og over alle de andre provinsene i Canada når det gjaldt leseferdigheter (se figur 8 vedlegg 1). I naturfag skåret Alberta høyere enn alle de andre Canadiske provinsene og lå rett under Finland, Japan og Kina. I

problemløsningsstrategier var Alberta nummer fem av alle OECD landene og lå over det Canadiske gjennomsnittet (Bussiere, Cartwright og Knighton 2004).

De norske 15 åringene lå under gjennomsnittet på flere områder i PISA undersøkelsen 2003. Elevene skåret omtrent som gjennomsnittet i OECD landene i matematikk og de skåret betydelig lavere i naturfag enn gjennomsnittet. Når det gjaldt selvregulert læring rapporterte norske elever om mindre vektlegging av læringsstrategier når de arbeider med matematikk, enn gjennomsnittet i OECD. Videre oppga norske elever lavere interesse for matematikk enn gjennomsnittet i OECD. Dette gjaldt særlig ferdighetstrening, utdyping og spesielt kontrollstrategier i matematikk. Når det gjaldt problemløsning lå de under gjennomsnittet (Kjærnsli 2004). Med andre ord er det en stor forskjell i resultatene til 15 åringer i Alberta, sammenlignet med 15 åringer i Norge.

Norges plassering i PISA har vært en underliggende årsak til at man har et ønske om å nå et ”kunnskapsløft” i den norske skolen. Et middel til å nå dette målet er å innføre et kvalitetsvurderingssystem, som skal ansvarliggjøre rektorer, lærere og elever i større grad enn før ved hjelp av blant annet vurdering. Det har vært mye debatt rundt dette, og man har vært opptatt av de negative konsekvensene vurdering og vurderingskriterier kan føre til. Utdanningsdirektoratet er i gang med å utforme nasjonale vurderingskriterier for elever også på barnetrinnet. Disse kriteriene har ennå ikke blitt vedtatt, men har likevel i flere sammenhenger ført til mye debatt og fått negativ omtale i mediene (jf s. 19). Utdanningsdirektoratet (2006) er klar over både positive og negative konsekvenser og omtaler disse på sine hjemmeside⁹. Noen av de positive sidene med å innføre slike kriterier er ifølge Utdanningsdirektoratet at de skal bidra til en likeverdig og rettferdig vurdering av elevene, styrke vurdering av den enkelte elev underveis gjennom presise faglige tilbakemeldinger og sikre sammenheng mellom underveisvurdering og sluttvurdering. Når det gjelder motargumenter og dilemmaer er Utdanningsdirektoratet klar over at nasjonale vurderingskriterier kan føre til en konkurrerende læreplan dersom kriteriene bare vektlegger noen av kompetansemålene. Videre kan det føre til en skjult innføring av karakterer på barnetrinnet og en mistillitserklæring til lærerprofesjonaliteten. Ved å vise til de to ulike skolene med ulik vurderingspraksis og læringsfokus i Edmonton (skole A og B), vil jeg avslutningsvis i min oppgave knytte de overnevnte positive og negative konsekvensene til

⁹ http://www.utdanningsdirektoratet.no/upload/Brev/Elevvurdering_i_Kunnskapsloftet.pdf

funn fra min empiriske undersøkelse. Målet med dette er å se hvordan erfaringene fra Edmonton kan være med på å berike den norske vurderingsdebatten.

Jeg ser i etterkant at det kunne vært interessant for denne oppgaven om jeg hadde tatt for meg lærerutdanningen både i Norge og Alberta i større grad enn jeg har gjort. Årsaken til det er at lærerutdanningen i Alberta og Norge er veldig forskjellig og det er her lærernes profesjonelle grunnlag skapes. Grunnet begrenset tid i Edmonton og til denne oppgaven, valgte jeg å fokusere på læreres praksis i skolene fremfor deres utdanning. Likevel har lærerutdanningen kommet inn i flere tilfeller fordi informantene i Alberta har tatt samme utdanning. Jeg har heller ikke valgt å analysere innholdet i de provinsielle prøvene eller PISA undersøkelsens gyldighet og kvalitet. Det kunne vært interessant å sammenligne PISA undersøkelsens form og innhold med de provinsielle prøvene for å se om det var noen likhetstrekk mellom disse. Grunnet oppgavens begrensninger og omfang, er dette noe jeg må overlate til videre forskning.

2 Hva er vurdering?

Som nevnt innledningsvis har vi i Norge mange ulike definisjoner på begrepene vurdering og evaluering og de brukes ofte om hverandre. I følge Gundem (1998) er det i Norge blitt vanlig å bytte ut det engelske låneordet ”evaluation” med begrepet vurdering. I følge Gundem brukes uttrykket vurdering som hovedterm både i grunnskolen og videregående opplæring, men også ”evaluering”, som på latin betyr ”sette verdi, være verdt”, dukker opp med samme mening (Gundem 1998). Fordi man i Norge ikke har laget et tydelig skille mellom disse begrepene, vil jeg starte med å ta utgangspunkt i de engelske begrepene ”assessment” og ”evaluation” sett fra et amerikansk / canadisk ståsted. Dersom ”assessment” og ”evaluation” oversettes til norsk, vil man kunne oversette begrepene til vurdering og evaluering.

2.1 “Assessment” versus “evaluation”

Det har vært mange forsøk på å forklare forskjellen mellom begrepene ”assessment” og ”evaluation”. Jeg vil her gjøre skillet klarere ved å ta utgangspunkt i noen ulike forklaringer på disse begrepene. Venezky karakteriserer ”evaluation” som en tilbakeskuende (”backward – looking”) bedømmelse av et læreprogram for å svare på spørsmålet om det har lyktes. ”Assessment” har i følge Venezky en mer fremadskuende (”forwardlooking”) karakter, med sikte på å besvare spørsmålet om hva som videre må gjøres ut fra de resultater som påviselig er oppnådd (Venezky (1974) i Lysne 1999: 307). Venezkys definisjoner kan sees i sammenheng med de termer som ble brukt av lærere og pedagoger i Alberta. ”Alberta Assessment Consortium” (AAC 2005:71) forklarer ”assessment” som: “a process of collecting information on student achievement and performance to improve student learning” Denne definisjonen er i tråd med Black og Wiliams (1998a) definisjon der vurdering er: ”all those activities undertaken by teachers, and by their students, which provide information to be used as feedback to modify the teaching and learning activities in which they are engaged” (Black og Wiliam 1998a: 2). ”Evaluation” forklarer AAC som: “Judgment regarding the quality, value, worth of respons” (AAC 2005:71). AAC’s definisjon på evaluering ligner også Mervin og Stakes definisjon: ”some degree of judgement regarding goodness or worth” (Mervin (1968) og Stake (1969) i Lysne 1999: 308). Ut fra disse definisjonene tolker jeg det dit hen at ”evaluation” er mer en bedømmelse av kvalitet og

verdi. "Assessment" derimot, er innhenting av informasjon gjort både av lærer og eleven selv (jf Black og Wiliam 1998a). Da jeg slo opp disse begrepene i en internasjonal ordbok på Internett fikk jeg to lignende forklaringer av begrepene "assessment" og "evaluation". "Assessment" var "the process of documenting, usually in measurable terms, knowledge, skills, attitudes and beliefs"¹⁰ og "evaluation" var "the systematic determination of merit, worth, and significance of something or someone"¹¹. Dette er i tråd med de øvrige definisjonene jeg viser til.

Ifølge Scriven (1967) kan evalueringens formål være å besvare ulike spørsmål om ulike enheter. Spørsmål av denne typen kan være: Hvor godt fungerer dette instrumentet sett ut fra visse kriterier? Fungerer dette instrumentet bedre enn et annet instrument? Er bruken av dette instrumentet verdt kostnadene? Med andre ord går målingen ut på å samle og sammenligne informasjon med målsettinger og et ønsket resultat. Når det gjelder evalueringens rolle i en undervisningskontekst kan den variere i stor grad. Den kan være en del av en lærerutdanning, læreplanutvikling eller for eksempel kan det være en del av et feltarbeid som er knyttet til forbedring av læringsteori (Scriven 1967).

På bakgrunn av de definisjonene av vurdering og evaluering som jeg har henvist til, vil jeg tro at det kan skape forvirring for mange dersom man tar med seg evalueringsbegrepet inn i klasserommet og forveksler det med vurderingsbegrepet. Jeg vil her henvide til et eksempel på at vurdering har blitt sett på som noe som skal bedømme eller måle verdi, i stedet for å bidra til læring. I Dagsavisen 5. Mai 2006, ble en eventuell innføring av nasjonale vurderingskriterier kritisert fordi det kunne føre til en "bedømming" av elevene helt ned til 2. klasse (Henmoe i Dagsavisen 05.05.06). For å unngå en slik mistolkning i denne oppgaven, vil jeg overlate evalueringsbegrepet til forskningsinstitusjoner og andre organisasjoner som har som oppdrag å skaffe informasjon om kvaliteten i skolen og skolesystemet. Dette er i henhold til Utdanningsetaten i Oslos definisjon der: "evaluering brukes om forskning og oppdrag levert av universitet og høyskolesektor"¹². Denne definisjonen kan også knyttes til Vedungs (1998) definisjon på evaluering, som har som formål å skaffe et helhetlig bilde av hvordan noe fungerer i en organisasjon eller system. Vedung forklarer evaluering ("utvärdering" på svensk) som: "en nogrann efterhandsbedömning av utfall, slutprestasjoner

¹⁰ <http://en.wikipedia.org/wiki/Assessment>

¹¹ <http://en.wikipedia.org/wiki/Evaluation>

¹² Fra "Vurdering, Åpenhet og oppfølging" (2005) av Utdanningsetaten i Oslo.

eller forvaltning i offentlig virksomhet, vilken avses spela en roll i praktiske beslutssituasjoner” (Vedung 1998:20).

I denne oppgaven avgrensar jeg meg fra evalueringsbegrepet ved å ta utgangspunkt i det engelske begrepet ”assessment” (vurdering) som jeg igjen vil dele inn i ”vurdering AV læring” og ”vurdering FOR læring”. Disse to vurderingsformene har to ulike formål: kartlegging og læring. Kartleggingen kan man gjøre underveis ved hjelp av formativ vurdering, eller på slutten av læringsprosessen ved hjelp av summativ vurdering. Jeg vil her gå nærmere inn på begrepene summativ og formativ vurdering.

2.2 Formativ og summativ vurdering

I lærebøker har man i mange sammenhenger delt vurderingens formål i to ved at man vurderer læringsprosess og læringsresultat. Man har kalt vurderingen underveis i læringsprosessen for formativ vurdering, underveisvurdering eller prosessvurdering. Vurderingen på slutten av læringsprosessen, har man kalt for summativ vurdering, sluttvurdering eller resultatvurdering. Disse begrepene brukes om hverandre i ulike sammenhenger og det blir her redegjort for de ulike begrepene.

Begrepene summativ og formativ kommer opprinnelig fra Scriven (1967). Han brukte disse begrepene i tilknytning til evaluering av læremiddelproduksjon og brukte evaluering som begrep. Formativ evaluering pekte på nødvendigheten av å vurdere et læremiddel under utviklingsprosessen, slik at det ferdige produktet kunne bli best mulig. Med andre ord er vurderingen prosessorientert (”forwardlooking”). Den summative evalueringen skulle føre til at læremiddelet kunne bli utstyrt med en form for varefakta til slutt. I følge Engelsen (2006) er disse to begrepene glidd over til den norske evalueringsterminologien, og vi bruker dem i tilknytning til elevvurdering. Hensikten med formativ vurdering er da å forbedre elevens kompetanse, integrert i læringsprosessen. Hensikten med den summative vurderingen er å gi en beskrivelse av elevens kunnskaps- og kompetanse nivå på et gitt tidspunkt (Engelsen 2006: 121). Et annet skille som knyttes til de samme formålene er prosessvurdering og resultatvurdering.

I følge Lysne inkluderer prosessvurdering både uformell og formell vurdering som fokuserer på progresjonen i en læreprosess. Uformell vurdering defineres som ”all vurdering, både i

skolen og ellers, som ikke gjøres i henhold til autorisert forskrift” og formell vurdering blir definert som ”standardiserte vurderingsordninger som foreskrives av myndighetene” (Lysne 1999: 10ff). I følge Lysne skal den prosessorienterte vurderingen være en gjennomgående del av undervisning og veiledning. Den har som primært formål å lede, korrigere og styrke læreprosessen underveis ved hjelp av løpende tilbakemelding om elevens effektivitet, faglig fremgang og mestring i forhold til mål. Resultatvurdering brukes også både i uformelle og formelle sammenhenger. Den har som primært formål å gi en vurdering av læreresultater som skal oppnås på ulike stadier i et læreprogram og mest inngående og konkluderende ved fullføringen av det (Lysne 1999). Prosessvurdering og resultatvurdering kan sees i sammenheng med begrepene formativ og summativ vurdering. Underveisvurdering og sluttvurdering, som nå brukes i forbindelse med Kunnskapsløftet, kan også sees i sammenheng med de overnevnte begreper. Underveisvurdering er vurdering underveis i læringsprosessen og sluttvurdering er vurdering ved læringsprosessens slutt.

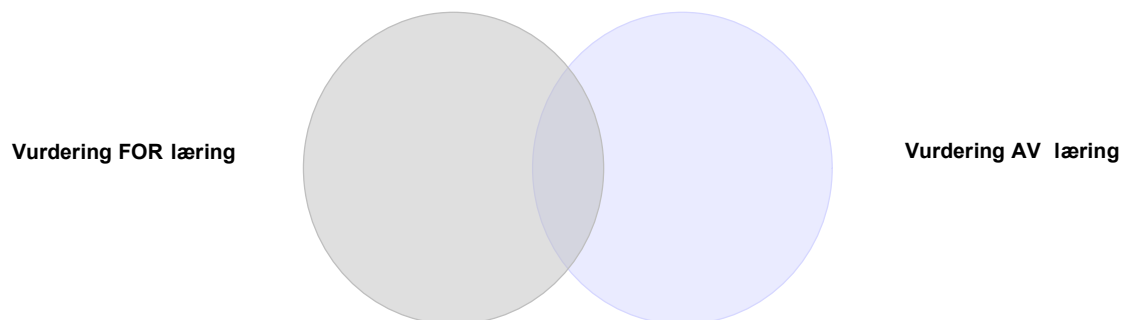
Formativ-, prosess- og underveisvurdering tar alle utgangspunkt i å vurdere læringsprosessen med formål om å forbedre læringsprosessen. Summativ-, resultat- og sluttvurdering har alle som formål å kartlegge et læringsresultat. Slik jeg tolker det, kan de sistnevnte begrepene sidestilles med ”vurdering AV læring”, mens begrepet ”vurdering FOR læring” innebærer all vurdering, kartlegging og innsamling av informasjon underveis i en læringsprosess, med formål om å forbedre læringsprosessen.

For å få en grundigere forståelse av skillet mellom vurdering FOR og AV læring, vil jeg ta for meg en oversikt over de to begrepene beskrevet av AAC (2005) (Se *figur 1* neste side). Dette skillet kan også beskrive skillet mellom formativ og summativ vurdering, men jeg vil ikke oversette vurdering FOR og AV læring direkte til disse begrepene. Årsaken til det er at ”vurdering AV læring” også kan brukes som en del av den formative vurderingsprosessen. Disse to vurderingsformene kan sees i sammenheng og dette forklares av AAC med en enkel modell (se *figur 2* neste side). Figuren bygger på Black og Wiliams forskning (1998a).

Vurdering FOR læring	Vurdering AV læring
<ul style="list-style-type: none"> - Læring kartlegges for å bestemme videre arbeid og for lage nye forslag til hva som skal gjøres - Undervisning og læring er uatskillige fra vurdering 	<ul style="list-style-type: none"> - Man sjekker hva som har blitt lært til dags dato
<ul style="list-style-type: none"> - Skal være en hjelp for lærere og elever til å forbedre læring 	<ul style="list-style-type: none"> - Har som også formål å gi informasjon til de som ikke direkte er involvert i den daglige læringsprosessen og undervisningen
<ul style="list-style-type: none"> - Tilbakemeldingen er spesifikk og deskriptiv og gis eleven ved hjelp av formell eller uformell kommunikasjon 	<ul style="list-style-type: none"> - Tilbakemeldingen gis eleven ved hjelp av tall eller karakterer
<ul style="list-style-type: none"> - Fokuserer på forbedring og sammenligner elevenes nåværende prestasjoner med hans eller hennes tidligere prestasjoner 	<ul style="list-style-type: none"> - Sammenligner en elevs prestasjoner med klassenivåets gjennomsnittresultater.
<ul style="list-style-type: none"> - Involverer eleven som en aktiv medspiller i læringsprosessen 	<ul style="list-style-type: none"> - Behøver ikke aktivt å involvere eleven som en partner i læringsprosessen
<ul style="list-style-type: none"> - Brukes sjelden for å sette karakter 	<ul style="list-style-type: none"> - Brukes ofte for å sette karakter
<ul style="list-style-type: none"> - Bruker pågående vurderinger som fokuserer på hele spekteret av læringsutbyttet – tenk på et fotoalbum 	<ul style="list-style-type: none"> - Bruker færre vurderinger som fokuserer på læringsutbyttet som en varig viktighet – tenk på et fotografi

Figur 1

Figur 2



Figur 2 er i henhold til Black og Wiliam (1998b), som mener at det ikke bør skilles i stor grad mellom disse to vurderingsformene. De mener det er urealistisk å forvente at lærere og elever kan praktisere et slikt skille i praksis. Utfordringen blir da å finne et mer positivt forhold mellom disse to formene for vurdering, slik at de summative testene faktisk måler deler av det som bli gjennomført i klasserommet. Black og Wiliam (1998b) tar opp noen former for hvordan dette kan gjøres: “1) Pupils should be engaged in a reflective review of the work they have done to enable them to plan their revision effectively 2) Pupils should be encouraged to set questions and mark answers to help them, both to understand the assessment process and to focus further efforts for improvement 3) Pupils should be encouraged through peer – and self assessment to apply criteria to help them understand how their work might be improved” (Black og Wiliam 1998b: 14). Hovedpoenget med disse tre punktene er at summative prøver kan være en positiv del av læringsprosessen. Ved å være aktivt med i prøveprosessen kan elevene se at de kan dra fordel av prøver fordi prøvene kan si noe om deres ståsted og hvor de er i sin læringsprosess (Black og Wiliam 1998b). Videre vil dette styrke de summative prøvenes validitet fordi det vil være en større sammenheng mellom de to vurderingsformene.

Fordi vi ikke har hatt formelle vurderingsformer på barnetrinnet, har det ennå ikke vært et tema i norske barneskoler å kombinere uformelle og formelle vurderingsstrategier (Lysne 1999). I stedet har man fokusert på den uformelle samtalen som virkemiddel for tilbakemelding. I følge Lysne (1999) ligger mulighetene for å få til et fruktbart samspill mellom lærer og elev, for å utvikle evnen til egenvurdering av egen ytelse og prestasjon, best til rette i den uformelle samtalevurderingen. Han tilføyer at dette også er fullt mulig i karaktervurderingen. Dette krever at karakteren forklares og drøftes med eleven i forhold til

prestasjonen og vurderingskriteriene. Han sier at det er avgjørende at eleven får hjelp til å skape seg en begrunnet forståelse av at det er logisk samsvar mellom prestasjonen og den karakter som er gitt. Dette er i samsvar med Black og Wiliams oppfordring.

I følge Lysne (1999) bruker de fleste land innen den vestlige kulturkrets uformell vurdering i barneskolen, noe som har endret seg med tiden. I Norge har vi fram til våren 2004, da nasjonale prøver ble forsøkt implementert, bare hatt uformell vurdering i barneskolen. I følge OECD (2005a) preges Norge av en såkalt "soft/dialogue-based supervision", uten klare standarder som illustrerer grad av ferdighetsoppnåelse i ulike årskull. Dette betyr at det kan være vanskelig for den enkelte lærer å vite om den enkelte elevs læringsprogresjon er god nok fordi de mangler klare standarder og vurderingskriterier. Videre kan de skape problemer for elevene fordi de ikke vet hva de skal jobbe mot og det blir vanskelig for både læreren og elevene selv å se progresjon i læringsforløpet. I følge Black og Wiliam (1998b) er det viktig å ha kriterier for å inkludere elevene mer i læringsprosessen og for å forstå hva som skal læres og hvorfor.

Jeg vil nå gå nærmere inn på forskningen til Black og Wiliam som er bakgrunnen for at "vurdering FOR læring" har fått en så stor oppslutning for lærere i Alberta og andre land. Eksempler på land som jobber med "vurdering FOR læring" for å motvirke et stort resultatfokus i skolen er England, Skottland, New Zealand og Australia (OECD 2005b).

2.3 Vurdering FOR læring i klasserommet

Black og Wiliams forskningsprosjekt kalt "Inside the black box – Raising standards through classroom assessment", var initiert av "British Educational Research Association". I følge Black og Wiliam (1998a: 1) var det behov for denne forskningen. De mente at tidligere politiske systemer i England har behandlet klasserommet som en "svart boks" fordi man har glemt å fokusere på hva som skjer på innsiden av klasserommets fire vegger. Departementet i England, har i følge Black og Wiliam de ti siste årene vært ivrige i å gjøre forandringer som kan bidra til å nå dette målet. Forbedringsprosjekter som blant annet nasjonale prøver, målsetting, initiativ til å forbedre planlegging i skolen og lederskap, mer rapportering og inspeksjoner er alle "means to the end". I følge Black og Wiliam fører ikke disse tiltakene til effektiv undervisning fordi noe mangler. Læring er drevet av det som skjer på innsiden av

klasserommet, og de mener at nivået kan bare øke dersom lærere kan takle denne oppgaven godt (Black og Wiliam 1998a).

Jeg vil her trekke en parallell til det som nå skjer i Norge. Vi vil heve nivået i skolen med den nye læreplanen "Kunnskapsløftet". Mange strategier for målig og vurdering og evaluering skal iverksettes for å heve kunnskapsnivået i den norske skolen. Man har fokusert på verktøy for å måle resultater, men verktøy for "vurdering FOR læring" har ikke vært i fokus på samme måte eller satt i system. Denne utviklingen tyder på at Black og Williams forskning også bør være relevant for norske klasserom. Black og Wiliam hadde tre spørsmål vedrørende formativ vurdering som de ville ha svar på i sitt forskningsprosjekt. Ved å gå igjennom 160 tidsskrifter ni år tilbake fant de til sammen 580 artikler om formativ vurdering. De tok i bruk 250 av disse kildene og publiserte en artikkel i tidsskriftet "Assessment in Education". Sammen med ledende eksperter fra Australia, Frankrike, Hong Kong, Sør Afrika og USA konkluderte de med at de kunne få positive svar på sine tre spørsmål.

Det første spørsmålet de tok for seg var: *Kan forbedring av formativ vurdering øke læringsutbyttet?* Forskningsartikler viste tydelig at det å forbedre metoder for formativ vurderingspraksis kan øke elevenes læringsutbytte og resultater. Dette viste også deres eget forskningsprosjekt som de gjorde i to skoler i England. Mange av studiene kom til den viktige konklusjonen at formative vurderingsmetoder hjelper spesielt de med lavt læringsutbytte og at dette ville øke det gjennomsnittlige læringsutbyttet i en klasse (Black og Wiliam 1998a).

På det andre spørsmålet: *Er det rom for forbedring av formativ vurdering i skolen?* Fikk de også positive svar. Black og William mener at hverdagsvurderingen kommer til kort på grunn av tre hovedproblemer. Det første hovedproblemet kaller de *ineffektiv læring*. Det vil si at det er en mismatch mellom læringsutbytte og vurderingspraksis. Noen årsaker til at dette oppstår, er at tester og prøver en tendens til å fremme utenatlæring og overfladisk læring selv om lærerne sier at de ønsker å utvikle forståelse. Mange lærere virket uvitende når det gjaldt denne inkonsistensen. Videre blir spørsmål og andre metoder som blir brukt ikke diskutert eller delt blant lærerne i den samme skolen. De blir heller ikke kritisk gjennomgått i forhold til hva de egentlig skal vurdere. Til sist er det vanlig for lærere i barneskolen at man vektlegger kvantitet og presentasjon av arbeidet og overser kvaliteten i relasjon til læring. Det andre hovedproblemet kaller Black og Wiliam *vurderingens negative*

påvirkning. Denne negative påvirkningen skyldes blant annet at man har overfokuserert på det å sette karakterer og på funksjonen dette har. Videre har det å sammenligne elever mer blitt sett på som å fremme konkurranse, enn å bli sett på som personlig fremgang. Som konsekvens vil vurdering gi elever med lav prestasjonsevne en følelse av at de ikke mesterer og en følelse av at de ikke er mottakelige for læring. Dette har gjort til at flere lærere har vegret seg for å vurdere elevene. Det tredje problemområdet kaller Black og Wiliam *vurderingens organisatoriske rolle*. Det vil si at i noen tilfeller kan læreres tilbakemelding til elevene ofte ha sosiale og organisatoriske funksjoner som ofte kan gå på bekostning av læringsfunksjonene. Videre kan lærere ofte forutse elevenes resultater på eksterne tester fordi deres egne tester overestimerer dem, samtidig som de vet for lite om deres elevers læringsbehov. Et annet viktig moment, er at innsamlingen av karakterer ofte kan bli mer prioritert enn det å analysere elevenes arbeid for å få informasjon om elevenes læringsbehov. For eksempel er det mange lærere som ikke tar i bruk vurderingsinformasjon gjort av andre lærere. Black og Wiliam tilføyer at problemområdene over ikke er noe som finnes i alle klasserom, men er eksempler på hvorfor det er behov for forbedring når det gjelder formative vurderingsmetoder. Beskrivelsene over bygger på forskning fra flere land som bygger på intervjuer, observasjoner og spørreundersøkelser i mange skoler (Black og Wiliam 1998a).

Det siste og kanskje mest interessante spørsmålet for lærere er dette: *Finnes det metoder for å forbedre praksis for formativ vurdering?* Svaret på dette spørsmålet var også ja, og Black og Wiliam (1998a) vektlegger spesielt tre hovedpunkter som må taes i betraktning for å oppnå slike metoder. Disse er: *elevenes selvtillit*, *egenvurdering* og utvikling av *effektiv undervisning*. Når det gjelder elevenes *selvtillit*, beskriver Black og Wiliam to syn på vurderingens effekt. Det negative aspektet er at en klasseromkultur som fokuserer mye på “gullstjerner”, karakterer og “ranking” i klassen, kan føre til at elevene ser etter den letteste utveien til å få den beste karakteren i stedet for å ta hensyn til sine læringsbehov. Elevene er mer opptatte av å finne det riktige svaret, enn å finne den riktige metoden. I stedet bør man ta tak i en klasseromkultur som er basert på en tro om at alle kan prestere og ha suksess. Her kan formativ vurdering være et godt hjelpemiddel hvis det brukes riktig. De beskriver dette positive aspektet slik: “Feedback to any pupil should be about the particular qualities of his or her work, with advice on what she or he can do to improve, and should avoid comparison with other pupils” (Black og Wiliam 1998a: 9). Når det gjelder punktet om *egenvurdering*, vektlegger Black og Wiliam viktigheten av at elevene selv har et bilde av hva de kan oppnå.

De refererer til Sadler (1998) som sier at når noen skal lære, består tilbakemeldingen de får av tre elementer: 1) et ønsket mål, 2) bevis på deres nåværende ståsted og 3) en forståelse av hvordan lukke gapet mellom de to. Alle disse tre må da bli forstått før man kan sette i gang en handling som skal forbedre læring. De oppsummerer dette slik: “for formative assessment to be productive, pupils should be trained in self – assessment so that they can understand the main purpose of their learning and thereby grasp what they need to achieve” (Black og Wiliam 1998a: 10).

Det siste punktet som Black og Wiliam tar opp, handler om planlegging av *effektiv undervisning*. Dette betyr at oppgaver må bli begrunnet av læringsmål og de kan bare fungere godt dersom elevenes forståelse er integrert i undervisningsplanleggingen. Dette oppsummerer de slik: “opportunities for pupils to express their understanding should be designed into any piece of teaching, for this will initiate the interaction whereby formative assessment aids learning” (Black og Wiliam 1998a).

I følge Black og William (1998a) er det innlysende at lærere må ha kunnskap om elevenes progresjon og læringsvansker, slik at de kan tilpasse sin undervisning til elevenes ulike behov. Lærerne kan skaffe seg informasjon og det de trenger på ulike måter ved hjelp av observasjoner, diskusjoner i klassen og skriftlig arbeid som elevene har gjort. Dersom eleven ikke blir gjort bevisste på hva de mestrer og hvordan de kan bli bedre i å bruke læringsstrategier, vil ikke deres refleksjon rundt egen læring bli optimal (Black og William 1998a). Black og Wiliam (1998a) mener at deres tre spørsmål bør være av stor betydning og at dette bør være noe utdanningsdepartementer og lærerprofesjonen bør studere nøye dersom de ønsker å øke elevenes læringsutbytte.

2.4 Vurderingskriterier

Som Black og Wiliam var inne på kjennetegnes effektiv undervisning av at læringsmål blir begrunnet. Vurderingskriterier spiller derfor en viktig rolle i forhold til dette. For å vurdere læringsprosess eller resultater er det viktig å sette kriterier for hva som må gjøres for å nå et læringsmål. Læringsmål kan være læreplanmål, eller delmål som læreren selv har utarbeidet (Gregory, Cameron & Davies 2002). I Norge har man ikke nasjonale vurderingskriterier, men det pågår en debatt i disse dager om hvorvidt vi bør innføre slike vurderingskriterier

eller ikke (jf del 1.2). Mange er skeptiske til konsekvensene dette kan føre til og jeg vil her ta for meg noen positive sider ved å bruke vurderingskriterier sammen med elevene.

Gregory, Cameron & Davies (2002) vektlegger at lærere kan sette kriterier for deres elever, de kan sette kriterier med deres elever og elever kan sette kriterier selv eller forhandle seg til dem. Dette er viktig i følge Gregory, Cameron & Davies (2002), fordi de har selv sett som lærere at jo mer elevene selv tar del i utviklingen av kriterier, jo bedre forstår de hva som er forventet av dem. Ved å involvere elevene i de kriteriene som allerede er satt og ved å sette nye kriterier i fellesskap, får elevene et eierskap til det de lærer og en større forståelse av hva de kan og hva de skal kunne (Gregory, Cameron & Davies 2002). De mener ikke at det skal settes kriterier for alt og at det er viktig å minne seg om målet med å bruke kriterier ved å stille seg spørsmålet: hvordan kan dette støtte elevens læring? I følge Alberta Assessment Consortium defineres kriterie som: "What students need to do to show they have achieved the learner outcomes" (AAC 2005:71). Med andre ord er målet med kriteriesetting å inkludere eleven i læringsprosessen.

I Alberta er vurderingskriterier en viktig del av læreplanen¹³ i alle fag og på alle nivå. Fagene deles opp etter hvert klassenivå og det settes kriterier og mål for hva elevene skal lære. Det er lagt frem mange praktiske eksempler på hvordan målene kan nåes og lærerne kan ta utgangspunkt i disse eller bruke egne metoder. På "Alberta Educations" hjemmeside¹⁴ står det: "Alberta's Kindergarten to Grade 12 curriculum is outlined in Programs of Study that identify what students are expected to learn and be able to do in all subjects and grades". Med andre ord er deres læreplan målrelatert i stor grad, og den setter standarden for hva som skal læres av alle elever i alle skoler i Alberta. Til tross for en stor grad av målrelatering gir læreplanen stor frihet i valg av metoder for å nå målene. Kriterier settes av lærere og elever både på individ -, klasse - og skolenivå og er med på å bevisstgjøre elevene på hvor de er i læringsprosessen.

I Norge har det også lenge vært læreplanen som har gitt føringer for hva elevene skal kunne. I L97 beskrives mål og hovedmomenter for hvert fag, men likevel er ikke kriteriebegrepet brukt. Hvilke vurderingskriterier som knyttes til hvert enkelt læringsmål er heller ikke beskrevet (jf L 97). Kriterier som blir laget på hver enkelt skole kan variere i stor grad, noe som da fører til at innholdet i samme læreplan blir vurdert ulikt fra skole til skole. Det kan se

¹³ læreplanen i Alberta heter "the Program of Studies"

¹⁴ http://education.gov.ab.ca/k_12/curriculum/bySubject/

ut som man i Norge har vegret seg for å bruke målrelatert vurdering i barneskolen, noe som kan ha en sammenheng med at man har skilt mellom formell og uformell vurdering. I en veiledning til L97 om elevvurdering (1998) uttrykkes en reservasjon overfor målrelatert vurdering: "Somme elever kan oppleve vurderinga som eit uttrykk for at dei kjem til kort i forhold til felles mål og kriterium". Det advares særlig mot hyppig bruk av ensidig vurdering på grunnlag av felles kriterier i forhold til elever som strever med å nå fellesmålene (KUF, 1998: 8). På barneskolen har man nedtonet den målrelaterte vurderingen av hensyn til "svake" elever. Likevel følger karakterene på ungdomsskolen et krav til målrelatert vurdering. Overgangen til ungdomsskolen kan da føre til store problemer for de "svake" elevene (jf Dale og Wærness 2006). Dette blir illustrert med denne situasjonen: "Vi har heile tida trudd at Mona var flink på skolen. Læraren hennar på barneskolen har heile tida fortalt oss kor greitt det går med ho (...). Heilt til ho tok til på ungdomsskolen. No blir ho stadig motlaus over dårlege karakterar, og vi skjønner ikkje at ho kan ha endra seg så mykje" (KUF, 1998: 9). Dette er et eksempel på at en brå overgang fra individrelatert vurdering, der målene settes ut fra elevens forutsetninger, til målrelatert vurdering kan være vanskelig, spesielt for de svake elevenes motivasjon. En innføring av målrelatert vurdering ved hjelp av kriterier i barneskolen kan derimot føre til at denne overgangen ikke blir så vanskelig for mange elever. Dette er i henhold til Black og Wiliam som mener at det er spesielt viktig at de svakeste elevene blir inkludert i læringsprosessen ved hjelp av mål og kriterier (Black og Wiliam 1998b).

For at kriteriene skal kunne være tydelige for elevene bør de også være tydelige for lærerne. Da jeg spurte de seks norske lærerne jeg intervjuet om hvem som avgjorde utviklingen av kriterier for elevvurderingen i deres skole og fikk jeg mange ulike svar. Det var en tydelig usikkerhet i forhold til hvem som satte kriterier fordi læreplanen ikke ga tydelige føringer for dette. Her er et eksempel på en uttalelse fra en kvinnelig lærer (F3) i skole F: *Kriterier ja ehh.. ja det kommer selvfølgelig litt an på hva du mener med akkurat det ordet, men det er klart at du ser jo litt sånn på læreplanen og du vet jo hva de skal igjennom fra trinn til trinn.* Jeg vil ikke bruke denne uttalelsen til å generalisere, men til å belyse hvordan en norsk lærer oppfatter kriteriebegrepet. Ingen av lærerne i Edmonton som jeg var i kontakt med var usikre da jeg stilte det samme spørsmålet. Det var klart for dem alle at det var læreplanen som først og fremst satte kriterier for elevvurderingen, men at de hadde frihet til å variere metodene.

2.5 Vurdering som profesjonsfaglig kompetanse

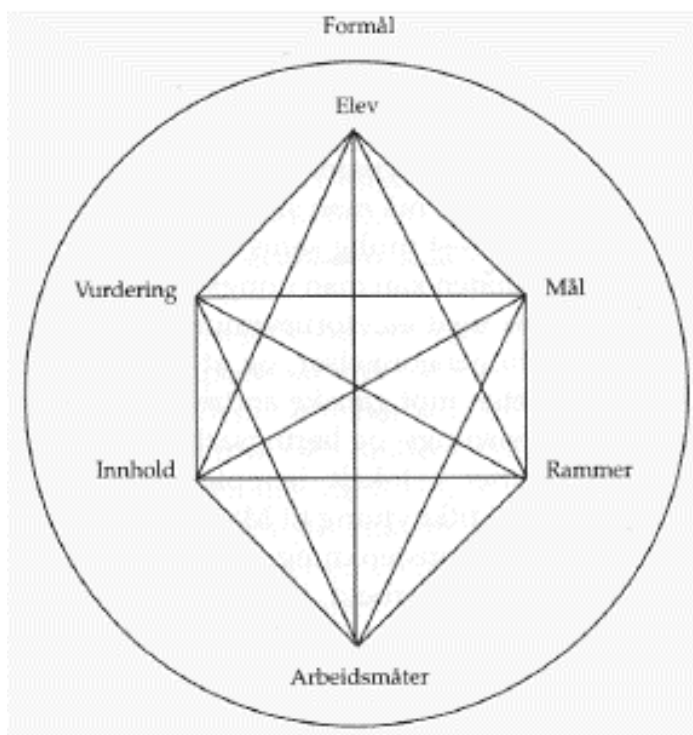
Begrepet profesjonsfaglig kompetanse har jeg utarbeidet på bakgrunn av norske fagtermer og prosjektet ”Vurdering som profesjonsfaglig kompetanse” som ble nevnt innledningsvis. Årsaken til at jeg bruker termer preget av den nordeuropeiske didaktikktradisjonen og ikke den amerikanske læreplantradisjonen, er ganske enkelt fordi didaktikk som begrep, ikke har en viktig rolle i USA eller i Alberta. Didaktikk forstås tradisjonelt med læreplan og undervisningsteori (Dale 1999: 18). Det vil si teori om hva som skal læres, hvordan innholdet skal organiseres og svar på hvorfor, det vil si å legitimere mål, innhold, metode og vurdering (Engelsen 2006). Jeg vil komme nærmere inn på dette begrepet litt senere i denne delen. I løpet av de fem månedene jeg tilbrakte på Universitetet i Alberta sammen med professorer, lærere og rektorer, ble ikke didaktikk som begrep nevnt en eneste gang. Da jeg spurte om noen visste hva begrepet innebar, var det bare et par professorer som hadde hørt om det, men det var ikke et begrep de brukte mye. Av den grunn ble det naturlig for meg å definere profesjonsfaglig kompetanse ut fra norske termer og begreper.

Begrepet profesjonsfaglig kompetanse har jeg først og fremst hentet fra prosjektet ”Vurdering som profesjonsfaglig kompetanse. Formålet med dette prosjektet var i følge Haugstveit å undersøke læreres praktiske vurderingsferdighet. De ønsket å gi innsikt i hvordan vurderingskompetansen blir forstått av et utvalg norske lærere på 5., 6. og 7. trinn. Lærerne i prosjektet ser på vurdering som en viktig del av profesjonskompetansen og de savner mer kompetanse på dette området (Haugstveit 2005). I følge Hoel (2005) er praksisnærhet, faglighet, tverrfaglighet og metodisk mangfold nøkkelbegrep som knyttes til profesjonsbegrepet i KUPP prosjektet. Med andre ord innebærer vurdering som profesjonsfaglig kompetanse at man har kompetanse til å bruke vurdering praktisk, faglig og metodisk. Hoels (2005) forståelse av profesjonsbegrepet har flere fellestrekk med Dales (1999) forståelse av begrepet ”profesjonalitet”. Dale vektlegger fire kunnskapselementer som utgjør lærerens profesjonsgrunnlag, og jeg vil her redegjøre for disse.

For det første er det bruk for kunnskap i de skolefaglige aktiviteter som læreren underviser i. For det andre har læreren bruk for fagdidaktikk der de skolefaglige aktivitetene settes inn i en intern fagplan. Fagdidaktikken omfatter spesifisering av mål, innhold, metode og vurdering i forhold til de enkelte skolefaglige aktiviteter. For det tredje er det i følge Dale viktig at lærerne kan opptre som kolleger med en innbyrdes kommunikasjon. Kolleger må

derfor kunne fortolke og analysere den generelle læreplanen som grunnlag for de skolefag som de underviser i. Det fjerde kunnskapselementet, er det overordnede grunnlaget for å gi læreryrket den profesjonelle identitet: pedagogikken. Pedagogikkens oppgave er å utvikle kriterier for kvalitet i utdannelsen. I følge Dale blir læreren profesjonell når pedagogikken blir kriteriegrunnlaget for å vurdere kvaliteten i elevenes læreprosesser. Med andre ord skal den profesjonelle skole utvikles med pedagogikken som grunnlag (Dale 1999:14). På bakgrunn av Dales beskrivelser, tolker jeg det dit hen at lærerens profesjonsgrunnlag består av de fire kunnskapsmomentene: kunnskap, fagdidaktikk, kollegialitet og pedagogiske kriterier for kvalitet. Jeg vil komme tilbake til disse punktene i analysen (jf del 6).

Tanken om den profesjonelle og autonome læreren har i mange år ligget til grunn for den didaktiske relasjonsmodellen til Bjarne Bjørndal og Sigmund Lieberg (1978 i Engelsen 2006:47). Modellen er blitt videreutviklet av blant annet Engelsen (2006), ved at alle de didaktiske faktorene skal sees i lys av et formål. Denne modellen bygger på og søker å videreutvikle den måten man mener lærere bør tenke på i en undervisningssituasjon og når de utformer undervisningen. Grunntanken med modellen er at man kan begynne med hvilken faktor man ønsker og at valget av en faktor berører de øvrige (Engelsen 2006).



Figur 3: Den Didaktisk Relasjonsmodellen (i Engelsen 2006:47).

Det er et gjensidig avhengighetsforhold mellom alle faktorene i denne modellen der ingen faktor er årsak og ingen er virkning. Modellen omfatter i tillegg til mål, innhold, metoder og arbeidsmåter i undervisningen, også ytre faktorer som kan påvirke planleggingen. Disse kalles rammefaktorer og preger lærernes handlingsrom.

Broadfoot skiller mellom vurderingens tre roller som er: 1) vurdering som hjelp for lærer og elev, 2) vurderingens rolle i læreplanen og 3) vurdering som et politisk verktøy for kontroll (Broadfoot 1996). I dagens globale samfunn der den internasjonale konkurransen er stor, er det vanskelig å komme vekk fra at vurderingens politiske kontrollfunksjon er til stede i stor grad. Ut fra et slikt syn på vurdering vil vurderingsfaktoren i den didaktiske relasjonsmodellen styre de øvrige faktorene i stor grad. I del 6 vil jeg redegjøre for i hvilken grad vurderingsfaktoren styrer lærernes bruk av vurdering som profesjonsfaglig kompetanse og som verktøy for læring i skole A og B.

3 Metodekapittel

Empirisk forskning er basert på generelle vitenskapsteoretiske perspektiver på hva kunnskap er, og hvordan kunnskap genereres ved bruk av ulike empiriske metoder (Lund 2002). Fordi mitt vitenskapsteoretiske utgangspunkt preger valg av metoder, vil jeg starte med å plassere meg vitenskapsteoretisk.

3.1 Vitenskapsteoretisk ramme

I denne oppgaven har jeg valgt et sosialkonstruktivistisk og fortolkende utgangspunkt for min forskning. Denne tilnærmingen aksepterer at virkeligheten er sosialt konstruert og det som blir observert er verdiladet og påvirkes av forskeren. Sosialkonstruktivisme kan defineres på flere måter men i denne oppgaven vil jeg ta for meg tilnærmingen kun i korte trekk. Sosialkonstruktivisme er en tilnærming som tar utgangspunkt i at både læring og kunnskap må sees i relasjon til kultur, språk og det øvrige fellesskapet som et individ hører til i (Wellington 2002). Dette er i tråd med definisjonen på kultur som jeg nevnte innledningsvis, som ser på kultur som “a living, historical product (...)” (jf Van Maanen & Barley i Rubin & Rubin 1995:20). Ut fra et sosialkonstruktivistisk utgangspunkt er vi mennesker selv aktive i å bygge opp eller konstruere en forståelse av oss selv og verden (Nygård 2003). Den fortolkende forskeren har sin bakgrunn i den hermeneutiske tradisjon der meningstolkningen er det sentrale. Begrepene samtale og tekst er essensielle og det legges vekt på tolkerens forforståelse av tekstens og samtalens emne (Kvale 2001). Min tilnærming står i motsetning til den positivistiske og eksperimentelle tilnærmingen, som erkjenner objektive sannheter og verdifri kunnskap. I tråd med den positivistiske oppfatningen, brukes eksperimentelle metoder til å studere mennesker og fenomener i kontrollerte omgivelser for å oppnå kvantitative data (Wellington 2001). Ut fra et sosialkonstruktivistisk og fortolkende utgangspunkt er kvalitative metoder det mest vanlige å bruke (Wellington 2002). Jeg vil starte med å ta for meg forskjellen mellom kvantitativ og kvalitativ metode.

3.2 Kvantitativ og kvalitativ metode

Kvantitativ metode knyttes tradisjonelt til den vitenskapsteoretiske tradisjonen positivisme, der forskeren foretar objektive målinger. Metoden brukes ofte innenfor naturvitenskapen, der kvantifisering og måling av data er sentralt. Grovt sagt vil man innenfor kvantitativ forskning karakterisere data som uttrykt i tall- og/eller mengdetermer, hvor hensikten er å få en representativ oversikt og gjøre statistiske generaliseringer (Kvale 2001, Wellington 2002).

Kvalitative metoder kjennetegnes ved at man beskriver karaktertrekk og egenskaper ved fenomener og mennesker. Målet med en slik metode er å tolke erfaringer, opplevelser, meninger og motiver i lys av sosiale omstendigheter. Dataanalysen er avhengig av forskeren som gjør en fortolkende og induktiv analyse. Induktive metoder kjennetegnes av at man går fra del til helhet og trekker generelle slutninger på grunnlag av enkelttilfeller. Deduktive metoder derimot, tar utgangspunkt i helheten ved at man går fra det generelle til det partikulære (jf Wellington 2000). Kvalitative metoder kan gi oss en pekepinn på hvordan noe er ved å gi innblikk i flere subjektive sannheter. Fordi jeg i min empiriske forskning ønsket å få et innblikk i læreres refleksjoner rundt vurdering og deres vurderingspraksis, har jeg valgt en kvalitativ empirisk tilnærming.

3.3 Planlegging av tema, utvalg og metoder

For å planlegge mitt forskningsprosjekt har jeg tatt utgangspunkt i en forskningsmodell kalt ”the ideal model for research” (Wellington 2000):

1. *Identifisere et felt man vil undersøke* – vurdering
2. *Velge tema* – Vurdering som profesjonsfaglig kompetanse og verktøy for læring i Edmonton, Alberta
3. *Bestemme tilnærming* – Dra tilbake til Edmonton for å gjøre kvalitative undersøkelser i to barneskoler. Gjøre en pilotundersøkelse i to barneskoler i Oslo
4. *Lage en plan* – lage intervjuguide, ta kontakt med skoler og andre informanter
5. *Innsamle informasjon* – dette ble gjort i hovedsak ved hjelp av observasjoner, intervjuer og innsamling av skriftlig dokumentasjon om vurdering i Alberta
6. *Analysere data* – analysere intervjuer
7. *Presentasjon av funn* – dette gjøres i analysekapittelet og i avslutningskapittelet.

Når det gjelder mitt utvalg av skoler og lærere har jeg gjort følgende valg. I Edmonton besøkte jeg til sammen fem barneskoler. Skole A og B valgte jeg på bakgrunn av informasjon fra to medstudenter som jeg gikk sammen med på et masterkurs ved universitetet i Alberta våren 2005. Før jeg startet prosjektet skrev jeg en mail til ca 20 medstudenter fra universitetet i Alberta som alle var rektorer og lærere. Jeg skrev at jeg ønsket å besøke to barneskoler. Jeg fikk to svar tilbake fra en lærer på skole A og en rektor på skole B, som begge ønsket at deres skole skulle delta i prosjektet. Jeg hadde sett for meg at skolene var ganske like, men etter å ha tilbrakt tre dager på hver skole opplevde jeg at skolene var veldig forskjellige.

Den første skolen (A) jeg gjør rede for var en "charterskole". Det vil si en offentlig skole som setter visse krav til elevene som kommer inn. De har en egen filosofi som de bygger sin undervisning på (jf del 5.1). Denne skolen godtok bare elever som presterte over et visst nivå og de hadde heller ingen elever med spesielle behov og lærevansker. De hadde et undervisningsopplegg som alle elever skulle ta del i som de kalte "whole group learning". Det akademiske nivået var høyt og det var et stort fokus på resultatbaserte prøver ("vurdering AV læring"). Undervisningen var tradisjonell og autoritær i stor grad. Skolen lå i et alminnelig middelklassestrøk og hadde en stor andel innvandrere fra Asia og India.

Den andre offentlige skolen (B), var styrt av EPSB og lå også i et vanlig middelklassestrøk nord i Edmonton. De hadde et program de kalte for "gifted class" som var for elever med spesielle behov for tilpasning. De hadde og et vanlig ordinært program uten spesielle tilpasninger. Denne skolen fokuserte på "vurdering FOR læring" på alle nivå og ønsket å tilpasse undervisningen best mulig til hver enkelt elev.

Jeg kom i kontakt med rektorer på skole C, D og E ved hjelp av et distriktsmøte med rektorer i regi av EPSB. Disse skolene ble også styrt av EPSB, men de hadde alle ulikt "instructional focus"¹⁵ og en ulik sammensetning av elever. Den tredje skolen (C) lå i et middelklassestrøk hvor det var mange innvandrere og hadde ingen spesiell inndeling. Den fjerde skolen (D) lå i sentrum nær universitetet i Alberta og hadde et samarbeid med universitet der forskere var involvert i undervisningsprogrammet i noen klasser. Den femte skolen (E) som lå i det samme strøket hadde nesten 45 % elever med behov for spesialundervisning.

¹⁵ Felles fokus som alle i skolen skal jobbe for

For å begrense mitt datamateriale har jeg valgt å intervju tre lærere på skole A og tre lærere på skole B, som alle underviste i 5. eller 6. klasse. Dette gjorde jeg for å gjøre mine funn mer sammenlignbare med funnene i prosjektet "Vurdering som profesjonsfaglig kompetanse" og fordi disse trinnene er preget av de provinsielle prøvene på 6. trinn. Årsaken til at jeg ikke har valgt 7. trinn, er at jeg ikke hadde tilgang til dette trinnet i den ene skolen i Edmonton. Med andre ord er mitt utvalg ikke helt tilfeldig.

Før jeg startet mine observasjoner og intervjuer i Edmonton, gjorde jeg en pilotstudie i to barneskoler i Oslo (skole F og skole G). Der utførte jeg intervjuer med 6 norske lærere på 5. og 6. trinn, med utgangspunkt i den samme intervjuguiden som ble brukt i Edmonton. Jeg tilbrakte tre dager i skole F og G, slik jeg gjorde i skole A og B. Dette gjorde jeg først og fremst for å utprøve mine metoder og for å bli bedre kjent med den norske vurderingskulturen. Jeg har ikke navngitt noen av mine intervjuinformanter og heller ikke skolene jeg har vært på. Dette har jeg gjort bevisst fordi jeg ikke ser det som vesentlig å spesifisere datamaterialet ved å bruke navn. På neste side har jeg laget en tabell der alle de 12 intervjuinformantene i Edmonton og Oslo er presentert. Oversikten tar utgangspunkt i 1) alder, 2) antall år erfaring som lærer og 3) utdanningssted. "U of A" er en forkortelse for "University of Alberta" og "HiO" er en forkortelse for "Høgskolen i Oslo".

EDMONTON			OSLO	
Skole A	Skole B		Skole F	Skole G
<u>A1: Kvinne</u>	<u>B1: Kvinne</u>		<u>F1: Kvinne</u>	<u>G1: Kvinne</u>
1. 60 år	1. 53 år		1. 36 år	1. 36 år
2. 12 års erfaring	2. 18 års erfaring		2. 11 års erfaring	2. 10 års erfaring
3. Utd: U of A	3. Utd: U of A		3. Utd: HiO	3. Utd: HiO
<u>A2: Kvinne</u>	<u>B2: Kvinne</u>		<u>F2: Mann</u>	<u>G2: Kvinne</u>
1. 28 år	1. 49 år		1. 33 år	1. 29 år
2. 7 års erfaring	2. 18 års erfaring		2. 7 års erfaring	2. 4 års erfaring
3. Utd: U of A	3. Utd: U of A		3. Utd: HiO	3. Utd: HiO

<u>A3: Mann</u>	<u>B3: Mann</u>	<u>F3: Mann</u>	<u>G3: Mann</u>
1. 55 år	1. 34år	1. 33 år	1. 31 år
2. 18 års erfaring	2. 1 års erfaring	2. 2 års erfaring	2. 4 års erfaring
3. Utd: U of A	3. Utd: U of A	3. Utd: HiO	3. Utd: HiO

Ved siden av å besøke fem barneskoler i Edmonton, fikk jeg i løpet av mars måned 2006 flere kontakter enn jeg hadde regnet med. Mine tidligere studievenner som både var lærere og rektorer og professorer ved "Department of Educational Policy" ved Universitetet i Alberta hjalp meg med å få tak i nøkkelpersoner innen vurderingsfeltet i Alberta. Informantene jeg snakket med var: 1) to professorer ved the University of Alberta og lærerutdanningen der, 2) leder for Alberta Assessment Consortium, 3) leder og to ansatte ved "Department of Assessment" i EPSB, 4) leder for "Department of Provincial Testing og Achievement" ved "Alberta Education"¹⁶ og 5) to ansatte ved "Alberta Teacher Association" (ATA). Disse informantene ga meg generell informasjon om vurderingens rolle i utdanningssystemet og de hjalp meg med å få en bred kunnskap om vurderingskulturen i Alberta og Edmonton. Jeg vil bruke skriftlig informasjon fra disse informantene som kilder i denne oppgaven. Videre har jeg selv skrevet notater fra mine observasjoner og mottatt kopier av skriftlig vurderingsmateriale fra lærere og rektorer på de ulike skolene.

3.4 Komparativ metode

I de senere år har det blitt en gradvis større politisk interesse for komparative studier som sammenligner likheter og forskjeller mellom ulike nasjonale strukturer. En årsak til dette er at man ønsker å finne løsninger på egne nasjonale problemer (jf Broadfoot mer fler 2003). Et eksempel på dette kan man se i forbindelse med diskusjonen rundt hva Finland gjør i sitt utdanningssystem, for å få så gode resultater som de fikk i PISA undersøkelsen 2003 (jf *figur 5* og *figur 6* i vedlegg 1). Finlands historie blir i flere sammenhenger trukket frem som en viktig faktor for hvordan lærerne praktiserer sin undervisning. En av Finlands mest kjente historikere Matti Klinge, sier til Aftenposten 12.12.04¹⁷ at det er måten Finland har møtt sine historiske utfordringer på som har gitt grunnlaget for deres skolesuksess. Han sier at mens

¹⁶ Utdanningsdepartementet i Alberta

Norge og andre vestlige land etter krigen mottok store summer i Marshallhjelp fra USA, var Finland nødt til å betale tunge krigserstatninger til Sovjetunionen. Dette var med på å tvinge Finland over i en meget rask industrialisering, der alle nasjonens ressurser måtte mobiliseres og skolen var et middel i denne prosessen. Alberta har som Finland gjennomgått en økonomisk krise og dette har satt konsekvenser for utdanningssystemet i disse landene. Jeg vil komme tilbake til hvordan dette har preget den politiske utviklingen i Alberta i del 4.

Flere steder i oppgaven sammenlignes Alberta med Norge. Det blir redegjort for vurderingskulturen i disse landene som er preget av både utdanningspolitikken og læreplantradisjonen de tilhører. Med andre ord sammenlignes data på både system – og individnivå. Dette er i tråd med Broadfoot med flere (2003), som mener at komparative studier kan gjøres på flere nivåer. Det vil si mellom individer i forskjellige klasserom, mellom individuelle klasser og skoler, mellom geografiske regioner og mellom nasjonale stater. Jeg tar utgangspunkt i deres beskrivelse av komparative studier som forklaring på hva komparative studier innebærer:

At one extreme they are uniquely well placed to study the impact of international phenomena such as multiculturalism, globalization and postcolonialism; at the other, they are equally well placed to study individual expectations and attitudes and the impact of peer group cultures and family life, all of which have an impact on educational experience. Comparative studies arguably have a unique power to reveal what that influence is and how it is realized in practice (Broadfoot med fler 2003:7).

Ut fra denne beskrivelsen er målet med komparativ metode blant annet å få frem relasjonen mellom kultur og læring (jf Broadfoot mer fler 2003). Dette er igjen i tråd med definisjonen på kultur, som ser kultur et historisk produkt som preger menneskers handlinger (jf Van Maanen & Barley i Rubin og Rubin 1995). Utdanningshistorien til Alberta, som ble omstrukturert på 1980 tallet, viser hvordan nyliberale tendenser og fokuset på ansvarliggjøring av ansatte i skolesektoren har påvirket skolesystemet (jf Kachur 1999). Hvordan den sterke vurderingskulturen har preget lærernes vurderingspraksis blir i denne oppgaven undersøkt ved hjelp av innsamling av empiri fra to skoler i Edmonton (A og B).

¹⁷ Se artikkel på: <http://forbruker.no/jobbogstudier/studier/article929687.ece>

3.5 Det kvalitative forskningsintervju

I følge Kvale (2001) er et forskningsintervju en konversasjon som har en viss struktur og hensikt. Videre er forskningsintervjuet ikke en konversasjon mellom likeverdige deltagere, ettersom det er forskeren som definerer og kontrollerer situasjonen. Temaet for intervjuet gis av intervjueren, som kritisk følger opp intervjupersonens svar på spørsmålene (Kvale 2001). Intervjuformen jeg brukte var semistrukturert. Det vil si en mellomting mellom strukturert intervju og ustrukturert intervju. Fordelen med strukturerte intervjuer er at de kan være lettere å analysere fordi det er mer forutsigbart. Dette gjør intervjuene i mindre grad fleksible. Den som intervjuer har mest kontroll og bestemmer gangen i intervjuet. Ustrukturerte intervjuer, derimot kjennetegnes av at de er veldig fleksible og uforutsigbare. Man har kontroll fra begge sider, den som blir intervjuet bestemmer gangen i intervjuet og det kan være vanskelig å analysere. Jeg visste ikke hvilke svar jeg ville få og måtte derfor tilpasse spørsmålene etter den konteksten jeg var i, og etter svarene jeg fikk. Det vil si at mine intervjuer var semistrukturerte fordi de ikke var helt forutsigbare, litt fleksible og jeg hadde mest kontroll som intervjuer (jf Wellington 2000:75). Målet med mine intervjuer har vært å få fram lærernes tanker og refleksjoner rundt vurdering, og få et inntrykk av hvordan vurdering brukes som profesjonsfaglig kompetanse og som verktøy for læring. Jeg kunne valgt å bruke spørreskjema med et større utvalg for å styrke validiteten¹⁸, men fordi jeg ville få tak i refleksjonene til intervjupersonene, ville det ikke vært hensiktsmessig å fremstille funnene i form av tall eller mengdestørrelser.

Før jeg laget intervjuguiden (se vedlegg 1), tok jeg kontakt med Bjarne Øygarden og Tove Brit Haugstveit fra prosjektet "Vurdering som profesjonsfaglig kompetanse" (2006). Jeg fikk tilgang til deres intervjuguide fra prosjektet og jeg brukte deler av denne som utgangspunkt for min intervjuguide både for norske og canadiske lærere. Jeg så det som hensiktsmessig å stille mange av de samme spørsmålene, fordi jeg hadde som hensikt å undersøke det samme temaet. Jeg lot mine intervjupersoner se på spørsmålene samtidig som jeg intervjuet dem. Årsaken til at jeg gjorde dette var fordi spørsmålene krevde refleksjon og lærerne trengte ofte litt tid på å svare.

For å minske muligheten for mistolkning under datainnsamlingen, tok jeg opp alle intervjusamtalene på digital diktafon, slik at jeg kunne transkribere intervjuene nøyaktig.

Dette gjorde også at det ble det lettere for meg å konsentrere meg om selve intervjusamtalen og ikke på notatskriving. Etter å ha tatt opp alle intervjusamtalene overførte jeg dem til min datamaskin. Jeg gjorde transkriberingen selv, ord for ord, hvor jeg tok med både mine spørsmål og deres svar. Ulempen med å bruke diktafon er at det fører til store mengder datamateriale og det tar tid å transkribere alle intervjuene (Wellington 2000: 86). Likevel føler jeg at dette var den beste måten å gjøre det på, fordi det ga meg muligheten til å lytte til samtalene flere ganger. I transkriberingen av mine intervjuer valgte jeg å understreke ord og setninger som var av betydning for min analyse. Kvale (2001) kaller dette meningsgenererende metoder, og hevder de kan få frem sammenhenger og strukturer som er av betydning for undersøkelsen.

I min analyse av hvordan de ulike lærerne brukte vurdering som profesjonsfaglig kompetanse tok jeg utgangspunkt i disse fire spørsmålene som jeg vil komme tilbake til i del 6. Da jeg gikk igjennom de ulike intervjuene, systematiserte jeg de ulike svarene fra informantene i fire bolker basert på de fire spørsmålene. Ettersom jeg leste igjennom intervjuene flere ganger oppsto det en ny analysekategori som jeg redegjør for i del 6.1.5 der jeg tar for meg de to skolenes ”skolekode”.

3.6 Observasjon

Jeg gjorde flere observasjoner i ulike klasser i de to barneskolene i Edmonton hvor mine intervjuinformanter jobbet. Deltagende observasjon vil si at forskeren mer eller mindre deltar i de handlingene som blir studert (jf Wellington 2000: 95). Ved å gå rundt å hjelpe elevene i timene, snakke med dem og gjøre observasjoner fikk jeg et innblikk i hvordan lærerne jobbet i praksis. Jeg gjorde dette i noen grad, men for det meste var det vanlig observasjon ved at jeg satt stille og observerte det som skjedde. Fordi jeg ikke selv er lærer, så jeg det som viktig å få et innblikk i lærerhverdagen både i Oslo og Edmonton før jeg startet mine intervjuer. Både i Oslo og Edmonton observerte jeg i hovedsak lærere som jobbet med 5. og 6.trinn i skole A, B, G og F. Men jeg tilbrakte også minimum 1 dag på hver av de øvrige skolene (C, D og E). Jeg observerte gjennomsnittlig lærerne i en klassetime hver. Det som var mest oppsiktsvekkende var at alle de fire offentlige skolene i Edmonton (B, C, D og E) synliggjorde vurderingskriterier og læringsmål på veggene i klasserommene. Dette gjorde de

¹⁸ Se definisjon s. 41

ikke på samme måte verken i skole A og eller i de to norske skolene. En slik synliggjøring av læringsmål og kriterier gjorde at jeg gjorde mange nyttige observasjoner av hva de fokuserte på i de ulike skolene. Ved siden av observasjoner og intervju samlet jeg også inn skriftlige dokumenter fra alle skolene jeg besøkte for å få et innblikk i skolens filosofi og ”instructional focus”. Jeg var også inne på deres Internett sider.

3.7 Validitet, reliabilitet og generaliserbarhet

Jeg vil avslutte denne delen med å ta for meg begrepene validitet, reliabilitet og generaliserbarhet som er viktig for å undersøke en metodes vitenskapelighet (Kvale 2001). Begrepet validitet er viktig i forskningsmetodisk sammenheng. Validitet referer til om en metode, en test eller et forskningsverktøy, faktisk måler det det skal måle (Wellington 2000). Jeg opplever at mine observasjoner, intervjuene og innsamling av skriftlige dokumenter ga meg kunnskap om dette.

Reliabilitet blir ofte brukt i forbindelse med forskningsmetoder. Reliabilitet referer til om hvorvidt en metode eller teknikk er konsistent ved at man kan forvente de samme resultatene hver gang metoden eller teknikken blir anvendt. Le Compte og Preissle (i Wellington 2000: 31) mener at ingen forsker som studerer den sosiale verden kan oppnå total reliabilitet. De beskriver begrepet slik: “The extent to which studies can be replicated. It assumes that a researcher using the same methods can obtain the same results as those of prior study. This poses an impossible task for any researcher studying naturalistic behaviour or unique phenomena”.

Noen øvrige problemområder i forhold til reliabilitet og generaliserbarhet, er beskrevet av Donald Schön (i Wellington 2000: 32). Han påpeker fire problemer innenfor det han kaller ”public learning”. For det første mener Schön at datainnsamling er en politisk prosess. Det er ikke all data man har tilgang på, og dette kan skape problemer. For det andre går data fort ut på dato og at det man samler inn kan bli fornyet før man har satt i gang analysen. Han påpeker også at det er problematisk at man fortolker data forskjellig ut fra ulike synspunkt og bakgrunnsforståelse. Det tredje punktet han påpeker, er problemet med å designe et hvilket som helst vitenskapelig eller systematisk eksperiment når man studerer et samfunn. Han mener at det nesten aldri er mulig å holde noen variabler konstante mens man manipulerer andre. Til sist tar han opp det han ser på som hovedproblemet som omhandler

generalisering ut fra resultater fra hvilket som helst studie. Schöns fire problematiseringer tar i hovedsak for seg områder for offentlig planlegging og politikk, men hans konklusjoner er relevante i forhold til det å samle inn data i utdanningsfeltet (Wellington 2000). Spesielt er punktet om tilgang et viktig punkt i mitt forskningsarbeid. Fordi jeg hadde lite bakgrunnskunnskap om utdanningssystemet i Alberta, var jeg avhengig av lokale informanter som kunne gi meg tilgang til skoler og andre nivåer i utdanningssystemet. Det andre punktet som omhandlet bakgrunnsforståelse er også viktig i min forskning. Min bakgrunn som norsk student og med erfaring fra et annet land og utdanningssystem kan ha ført til at jeg har tolket mine funn annerledes enn hva en canadisk student ville ha gjort. Dette kan også være positivt i et slikt studie der man studerer en annen kultur enn sin egen, fordi forskjellene kommer tydelig fram. Det er kjent for antropologer at man må ha en kulturell avstand mellom forsker og subjekt for å lære hva en kultur virkelig består i (Howell og Melhuus 2001). Videre er punktet som omhandler generalisering sentralt. Mitt materiale er altfor lite til å gi grunnlag for generaliseringer i forhold til vurderingspraksis i offentlige skoler i Edmonton i sin alminnelighet, og det har heller ikke vært mitt mål å undersøke dette. Likevel mener jeg at jeg har sett tendenser som er interessante og som kan tolkes for et hypotesegenererende formål.

4 Vurderingskulturen i Alberta

Grunnoppfølgingen i Alberta er et provinsielt og kommunalt anliggende og styres av provinsielle læreplaner, der bydistriktene er ansvarlige for at opplæringen skjer etter de provinsielle føringene. Det er 42 offentlige skoledistrikter i Alberta og 17 separate skole distrikter som alle er offentlig støttet. 16 av de 17 separate skole distriktene har et romersk-katolsk valgstyre og et distrikt har et protestantisk styre. Utenom de offentlige skolene har Alberta et utvalg av charterskoler, private skoler og hjemmeskoler. Alle disse skolene er autorisert, regulert og delvis støttet av "Alberta Education". Offentlige, separate skoler, charter skoler, private skoler og hjemmeundervisning følger læreplanen som er godkjent av "Alberta Education". "Alberta Education" administrerer også de provinsielle prøvene og avgangseksamener for elever på videregående og de sertifiserer lærere som skal jobbe i alle de ulike skolene i Alberta (ATA 2002). Jeg vil her ta for meg noen bakenforliggende årsaker til Albertas utdanningspolitiske utvikling som har påvirket resultatfokuset i Alberta.

4.1 Utdanningspolitiske trekk

Fram til 1960 tallet var Albertaprovinsen preget av en pedagogikk som var påvirket av John Dewey og progressive strømninger fra USA. Etter Sputniksjokket i 1957 og den kalde krigen endret de utdanningspolitiske føringene seg. En viktig årsak til disse endringene var den store økonomiske krisen og arbeidsløsheten i Alberta på slutten av 1970 tallet (Mazurek 1999). Utdanningssystemet ble sett på som en syndeby for dette og det ble behov for endringer i utdanningssystemet. Den progressive pedagogikken ble kritisert nasjonalt og et viktig bidrag til dette var Hilda Neathby's "So Little for the Mind" (i Mazurek 1999: 10f). Hun kritiserte den progressive pedagogikken ved å skrive i sin bok: "the standards of rigour had debased the most worthwhile subjects by treating them superficially and by including them as trivial studies, and had contributed to the decline of the best social, ethical, and political values to emerge out of western intellectual tradition. The cure was obvious: a return to a core of essential knowledge and disciplines and a reaffirmation of traditional values" (i Mazurek 1999: 10). Dette var starten på en omvending til nyliberale strategier i Canada og spesielt i Alberta.

I følge Harrison og Kachur (1999:xiii) var Alberta: “the first province to frontally attack the notion of equality of opportunity, replacing it with New Right policies redefining ”opportunity” according to a new logic – competition, effectiveness, and standards – and to emphasize choice, vocationalism and marketization”. Alberta var en av de provinsene og landene som hang seg på denne bølgen som Harrison og Kachur kaller nyliberalisme. For å forklare dette begrepet tar jeg utgangspunkt i David Harveys definisjon: “Neoliberalism is in the first instance a theory of political economic practices that proposes that human well – being can best be advanced by liberating individual entrepreneurial freedoms and skills within an institutional framework characterized by strong private property rights, free markets, and free trade. The role of the state is to create and preserve an institutional framework appropriate to such practices” (Harvey 2005: 2). Nyliberalismens første seier var i England under Margaret Thatcher i 1979 og i USA under Ronald Reagan i 1980. Den andre nyliberale bølgen slo til i 1984 i New Zealand og så i Canada gjennom Alberta (Harrison og Kachur 1999). Denne politiske retningen gjenspeilte seg i Albertas politiske føringer på begynnelsen av 1980- tallet. I Alberta ble det politisk ytret et ønske om større grad av ansvarliggjøring og rapportering av resultater, og nasjonale vurderings- og evalueringssystemer på alle nivåer i utdanningssystemet ble midler til å nå dette.

Allerede i 1982 ble det implementert provinsielle prøver i engelsk, matematikk, naturfag, og social studies på 6. 9. og 12. trinn på en roterende basis i Alberta. I 1992 fikk de også prøver for elever i 3. klasse. Det var mye motstand fra lærerne når det gjaldt disse prøvene for de mente at det truet deres profesjonalitet (ATA 2002). Likevel 23 år etter, brukes prøvene i alle skoler i Alberta til tross for at mange lærere fortsatt er i mot en slik resultatvurdering. Målet for disse prøvene er: 1) å se om elevene lærer det de forventes å lære, 2) rapportere til de som bor i Alberta hvor godt elevene har nådd de provinsielle standardene på en gitt tid i sin skolegang, 3) hjelpe skoler, skoleautoriteter og provinsen til å forbedre læring hos elevene, og 4) oppnå nøyaktighet og ansvarlighet¹⁹. I tillegg til disse provinsielle prøvene har de ulike bydistriktene lokale ordninger og prøver. Et eksempel på dette er EPSB, der alle elever siden 1995 har distriktsprøver hvert år, ved siden av de provinsielle prøvene. Disse kalles ”Highest Level of Achievement Tests” (HLAT) og er en prøve for lesing og skriving for alle klassenivåer. I EPSB er det vanlig at elevene starter i kindergarden fra de er ca 5 år gamle. Elevene får ikke karakterer før på våren i 1. klasse og får fram til da kommentarer. Tre

¹⁹ www.education.gov.ab.ca

ganger i året får elevene utlevert ”reportcards” som er en slags karakterutskrift der elevene får skriftlig dokumentert sine karakterer og kommentarer. I private skoler og charterskoler kan elevene få med karakterer fra de er fem år gamle. Med andre ord er barneskolene i Alberta preget av ”vurdering AV læring” i mye større grad enn barneskolene i Norge, der elevene ikke får formelle kommentarer.

4.2 Behov for “accountability” i Alberta

Som en konsekvens av nyliberalismen gjennomgikk læreplanen omfattende endringer i 1984 der mange valg for lærere ble innskrenket. Det ble innført obligatoriske kurs i karriere, personlig finans og ”life management”. Evaluering av lærere, elever og skoleprogrammer ble iverksatt. “Accountability” vil jeg forklare som: “the process through which individuals or organizations in the education system take responsibility for their actions and report on these actions to those who are entitled to the information” (ATA 2005:45). Dette begrepet kan på norsk oversettes til ”ansvarliggjøring” og blir brukt i flere sammenhenger for å forklare skolens ansvar. Fokuset på ”ansvarliggjøring” har blitt større også i Norge de siste fem årene etter at det i 2002 ble bestemt at vi skulle ha et kvalitetsvurderingssystem (jf del 7.1.2). Et eksempel som viser at ansvarliggjøring politisk blir satt på dagsorden i Norge, er kapittel 20 i NOU²⁰ rapporten ”I første rekke” (2003:16). Her blir skolens, skolelederens og skoleeiers ansvar beskrevet slik: ”Skolen har ansvaret for at elevene - som er eksistensberettigelsen for hele organisasjonen - skal tilegne seg, videreutvikle og skape en lang rekke kompetanser over et stort antall år (...) Skolen har også ansvaret for at læreren - som har som sin profesjon å legge til rette for andres læring - til enhver tid har en kompetanse som er tilstrekkelig god til å kunne gjøre en verdifull jobb”²¹

Omtrent 10 år før ”ansvarliggjøring” ble satt på dagsorden i Norge, ble dette prioritert i stor grad i Alberta i perioden 1993 til 1996. Jeg vil her redegjøre for fem dokumenter som viser Albertas utdanningspolitiske prioriteringer i denne perioden. Det første dokumentet ”Measuring Up”, gjorde rede for de mål hvert departement skulle nå, og skulle være en indikasjon på hvordan målene skulle nåes. Videre inneholdt reformen retningslinjer for hvordan hvert departement skulle bli evaluert i forhold til om målene var nådd og om politiske planer var iverksatt. Reformen ”Meeting the challenge”, spesifiserer hvordan

²⁰ Norges offentlige utredninger

²¹ http://odin.dep.no/kd/norsk/dok/andre_dok/nou/045001-020003/hov020-bn.html

Alberta vil møte disse utfordringene. "Roles and Responsibility in Education" tar for seg roller og ansvar til alle deltagerne i utdanningen. Der ble elever, foreldre, skoleråd, skoleetater, skoledirektører, rektorer, lærere, øvrige ansatte i skolen, distriktene og "Alberta Education" omtalt med roller og ansvar. "Teacher Policy reform" tok for seg kunnskap, ferdigheter og innstillinger som er forventet av lærere og som kreves for å sertifiseres. "Accountability in Education" fokuserte på rapportering på skole - og etatnivå. Denne femdelte businessplanen til departementet som gikk over tre år, førte til store endringer i utdanningssystemet. Noen av disse var blant annet: fokus på læreplanmål og standarder som gjorde skolen mer relevant, vurderingsprogrammer knyttet til læreplanmålene og standardene, samt evaluering av lærerutdanningen (Mazurek 1999).

Albertas utdanningspolitikk har ført til ulike konsekvenser. Til tross for sin lave arbeidsløshetsprosent (som nå er på 3,4 %, halvparten av landets gjennomsnitt), har landet fram til nå hatt landets laveste prosent når det gjelder å fullføre videregående skole. Bare 2/3 deler av videregående elever fullførte sin skolegang i 2003 – 2003²². Årsaken til dette er blant annet stor konkurranse og press i skolen, men også det faktum at det er lett å få godt lønnet arbeid uten endt skolegang. På grunn av provinsens gode oljeinntekter og god økonomi, har unge god tilgang til jobber der de kan tjene gode penger uten utdanning. Ranking av både skoler og elevresultater og det faktum at mange elever ikke fullfører videregående har vært følger av resultatfokuset i Alberta. Av den grunn har man iverksatt strategier for å motvirke dette. I den provinsielle læreplanen har man begynt å fokusere på formative strategier for å inkludere elevene i større grad i læringsprosessen og for å øke elevenes motivasjon (Jf OECD 2005b).

²² <http://www.united-online.ca/phpBB2/viewtopic.php?t=626&sid=81d27a5ee6ad46b33b2274918a7ee7b0>

5 Presentasjon av funn

I denne delen av min empiriske undersøkelse tar jeg utgangspunkt i en ”charterskole” og fire offentlige skoler som er en del av EPSB. Begrepet ”charter” skole kommer opprinnelig fra USA og har utviklet seg og spredd seg til Canada de siste 10 årene. Det er 12 charter skoler i Canada og alle ligger i Alberta. En charterskole kan forklares som: “A charter school is a school run by Directors. A locally elected Board governs the school. Further there is also an elected Parent Council that helps with the day to day activities, such as the hot lunch program and fund raising. A Charter school is a public school and therefore does not charge any tuition fees²³”. Charterskolen jeg besøkte er foreldrestyrt og følger sine egne retningslinjer, men den samme provinsielle læreplan. De har egne resultatbaserte prøver utover de provinsielle prøvene.

5.1 Skole A: Vurdering AV læring i fokus.

Denne ”charterskolen” ble etablert 16. mai 1996, og åpnet da sine dører for 280 elever fra 1. klasse til 8. klasse. I 1997 fikk skolen også 9. klasse og de har nå ca 440 elever. Skolen har ingen elever med behov for spesial undervisning og alle følger den samme undervisningen. For å sikre at elevene kan følge det høye faglige nivået som er på skolen, må elevene ta en kunnskapstest for å komme inn på skolen. Skolen tar imot elever som er over gjennomsnittet og alle skal følge det sammen undervisningsprogrammet som de kaller for ”whole group learning”. Skolen må som alle andre skoler i Alberta forholde seg til Albertas læreplan. Ut fra denne læreplanen har de utviklet sin egen læreplan som tydeliggjør læringsmålene i hvert fag på hvert nivå. Dette har vært gjort med mål om å finne et innhold som kan tilpasses skolens formål og for å utvikle en læreplan som møter de standarder som er satt for denne skolen.

I en rangering gjort på Internett, ligger denne skolen som nr. 30 av alle skolene i Alberta på grunn av elevenes høye resultater. Dette var noe både lærere og elever var veldig stolte av og denne oversikten hang på lærerværelset. På deres hjemmeside har de selv skrevet at: “the recent external evaluation compliments us on the strong culture of our school, which has come to be known as the "the way of this school". Med andre ord er det ikke vanskelig å se

at ”vurdering AV læring” er deres fokus. Deres skole blir sett på som en effektiv skole med høyt resultatnivå og for å nå høye resultater har de sin egen filosofi som de deler opp i fem punkter²⁴. Disse blir sett på som nødvendig for å oppnå ”achievement of excellence” i en effektiv skole: 1) Høye forventninger til både akademisk arbeid og oppførsel, 2) et behagelig skolemiljø med fokus på orden, 3) et sterkt fokus på instruksjon, 3) konsistent støtte fra hjem og skole, 4) en sterk oppfattelse av skolens forpliktelse og formål og 5) regelmessig vurdering av læringsutbyttet. Fordi skolen er styrt av the Company’s Act by a ”Board of Directors” som er alle foreldrene, er målene for denne skolen satt av foreldrene selv. De siste årene har det vært en stor pågang på denne skolen, spesielt fra foreldre med asiatisk og indisk bakgrunn. En årsak til det er at mange savner struktur, orden og resultatfokus i de offentlige skolene.

Etter å ha observert 4 ulike lærere i en tre dagers periode, satt jeg igjen med et inntrykk av at deres undervisningsform var tradisjonell i stor grad. Elevene satt for det meste på hver sin pult ikledd uniform. Lærerne ble kalt ved etternavn og det var stor grad av disiplin og respekt for de voksne. I timene gikk elevene opp til lærerne dersom de hadde spørsmål. Elevene fikk karakterer fra 5 årsalderen i ”kindergarden”, og jeg fikk beskjed om at de fleste elever i skolen lå på mellom 65 % og 80 %. Det var en stor åpenhet rundt karakterer, og klassene konkurrerte seg i mellom. Skolens vurderingsverktøy var først og fremst ”reportcards” som ble oppdatert tre ganger i året hvor lærerne samlet resultatene av det meste som elevene hadde gjort på skolen, i lekser og prøver. Dersom elevene manglet noen poeng kunne de gjøre bonusarbeid, der man kunne få opp til 10 % for å øke sin poengsum. Videre hadde alle elevene en hjemmeleksebok som foreldrene skulle sjekke hver dag for å se hvilke lekser de hadde. I forbindelse med at elevene får tilbake sitt ”reportcard” arrangeres det et ”foreldre – lærer intervju” der de går igjennom fag for fag og resultatene i de ulike emnene. Rektoren var veldig opptatt av gode resultater, og alle kriteriene for ”vurdering AV læring” som jeg henviste til i *figur 1* s. 22 var til stede i deres lærernes vurderingspraksis.

5.1.1 Observasjoner tatt fra en klasstime

Jeg observerte blant annet en naturfagstime, en fransktid og en matematikktid for både femte og sjetteklasse. Jeg vil her gi et eksempel fra en mattetid jeg observerte i en femteklasse. Denne klassen har som de andre klassene, en matteprøve i uken. Alle må levere

²³ Hentet fra ”Introduction to ... Charter School for prospective parents” (2006)

en ”kontrakt” hver uke der foreldrene har signert for å bekrefte at de hadde sett ukens gjøremål. Timen er lærerstyrt i stor grad og preget av ro og stillhet. Elevene var i gang med å rette en matteprøve da jeg var på besøk. Da de rettet prøvene, gikk jeg rundt og spurte om elevene likte matematikk. De smilte alle og sa ja de gjorde det. De hadde så vidt tid til å tittle opp og snakke med meg for de var så opptatt av å bli først ferdig med rettingen. Da jeg spurte en gutt om hvorfor han likte matte så godt svarte han: - *because it is challenging and I have to use my head a lot*. Denne gutten var motivert og gjorde det godt resultatmessig. Jeg får vite av læreren at de sammenligner resultatene til de fire matteklassene i grafer hver uke og at det er som en slags konkurranse som motiverer dem. Etter å ha rettet prøven blir de bedt om å gjøre noen individuelle oppgaver. Når de er ferdig skal de komme opp til læreren slik at hun kan sjekke om det er riktig. Dersom de har gjort alt riktig godkjenner læreren det skriftlig i boken. Dersom de har noe galt må de gjøre det på nytt og de får vite hva de har gjort galt og hva som er riktig i forhold til den strategien de bruker. Dersom de ikke får til en oppgave etter mange forsøk får de beskjed om at de må høre med en av ”ekspertene” som allerede har fått det til. ”Ekspertene” skal ikke gi riktig svar, men hjelpe dem med å velge riktig strategi. På slutten av timen går de igjennom gangetabellen og får beskjed om at dette må de vite til deres avsluttende eksamen.

5.1.2 Observasjoner tatt fra et foreldre – lærer- intervju.

Jeg vil her redegjøre for noen observasjoner jeg gjorde på et foreldremøte som var i forbindelse med at alle elevene skulle få tilbake sine ”reportcards”. Alle lærerne satt i en stor gymsal på hvert sitt bord og foreldrene gikk rundt til de ulike lærerne. Eleven kunne delta på konferansen, men behøvde ikke å være med. Jeg fikk lov til å lytte til samtalene til de ulike lærerne jeg hadde vært i kontakt med. Jeg fulgte 5 ulike lærere som alle gikk igjennom elevenes karakterer på hver prøve som var blitt gjort det siste halvåret. De ga karakterer i prosent og mange foreldre var veldig opptatt av hva de kunne gjøre for å øke prosenten til eleven. Eleven var involvert i samtalen i noen anledninger, men ikke ofte. Jeg deltok på en samtale der en elev som gikk i femteklasse fikk tilbakemelding fra sin fransklærer sammen med sin mor. Både moren og læreren var oppgitte fordi han hadde fått 0 % på flere prøver. På flesteparten av prøvene med 0 % hadde han glemt å skrive navnet sitt på prøven og det var et kriterium som måtte være med. Hvis navnet manglet, ville ikke læreren engang se på det han hadde gjort. Gutten satt med bøyd hode og lyttet til alle de lave

²⁴ Hentet fra ”Introduction to ... Charter school for prospective parents (2006)

prosentene som ble lest opp. I løpet av samtalen kom det fram at gutten hadde løyet for moren flere ganger og ikke gitt henne tilbake prøver som skulle godkjennes og underskrives. Læreren sa til dem begge at dersom dette ikke bedret seg måtte han nok etter hvert skifte skole fordi han ikke lenger nådde de standardene de hadde. - *Its hard out there*, sa både moren og læreren flere ganger for å understreke viktigheten av at han måtte bedre sine karakterer. Gutten var tydelig nervøs, og sa at han skulle forsøke å forbedre seg. På slutten av samtalen spurte jeg gutten om han likte å gå på skolen. Han svarte at han visste at han var heldig som fikk gå på denne skolen der lærerne var så dyktige. ”Men for å være helt ærlig”, sa han til slutt, ”jeg liker ikke å gå på skolen fordi det er alt for vanskelig”. Han sa at han hadde problemer med å følge med. Dette var et eksempel på en gutt som ikke passet inn i ”whole group learning” programmet som denne skolen hadde.

5.2 Skole B: Vurdering FOR læring i fokus

Skole B er en vanlig offentlig skole styrt av EPSB. Deres fokus er leseforståelse, noe som preger mye av deres arbeid. Her er et utdrag fra deres hjemmeside: “All students will improve their level of reading for meaning across all curriculum areas. We will measure growth through reading comprehension assessments, anecdotal observations, and standardized achievement tests”. Dette utdraget viser at de også er opptatt av ”vurdering AV læring”, men de bruker varierte vurderingsmetoder. Skolen har et differensiert program der noen klasser er inndelt etter nivå. Ved siden av det vanlige programmet har de et program som heter ”gifted” program der de jobber med tilpasset undervisning for elever over gjennomsnittsnivået. Dette innebærer at de gir ekstra undervisningsstøtte til elever som er på et usedvanlig høyt akademisk nivå. Elevene med lærevansker eller behov for spesialundervisning, er integrert i det normale programmet. Alt i alt har skolen 297 elever og skolebygningen er preget av både åpne og lukkede klasserom. Elevene sitter som regel i grupper og jobber sammen og de har ikke uniformer. Hele skolen er preget av at det henger arbeid på veggene som elevene har laget ved siden av læreplanmål og vurderingskriterier. Til og med på lærerværelset henger det kriterier på veggene for hva det vil si å være en profesjonell lærer. Vurderingsmateriell fra ”Alberta Assessment Consortium” er også hengt opp. Skolen ligger i et vanlig middelklassestrøk litt utenfor Edmonton sentrum. Elevene i ”kindergarden” får kommentarer og ikke karakterer, men elever fra første klasse får karakterer på slutten av året. Elevene på denne skolen blir også vurdert ut fra ”reportcards”, som de får utdelt tre ganger i året, men fokuset på å bruke prosenter er ikke så stort som i

skole A. De bruker også kommentarer ved siden av karakterene. De har ”studentled conference” og ”parent teacher conference” en gang i året. Rektoren er selv veldig involvert i ”vurdering FOR læring”. Videre var alle også kriteriene for ”vurdering FOR læring” som jeg henviste til i *figur 1* s. 22 til stede i lærernes vurderingspraksis.

5.2.1 Observasjoner tatt i en klasstime

På denne skolen observerte jeg tre ulike klasser: en vanlig 5. klasse, en vanlig 2. klasse og en 4. klasse som var ”gifted class”. Det er den sistnevnte timen jeg vil ta utgangspunkt i. Klasserommet var dekket med målsetninger og kriterier for hvordan målene kan nås. Elevene satt to og tre sammen i grupper. Timen var preget av ro og de sa ingenting uten å rekke opp hånden. De fremførte en del av et prosjekt de hadde jobbet med om leseforståelse. De sto en og en foran klassen og viste de andre elevene hvordan de har løst oppgaven og hvorfor de hadde gjort som de hadde gjort. Læreren ga hyppige tilbakemeldinger og elevene ble ofte bedt om å gi respons. De reiste ofte seg opp midt i timen og vandret rundt for å levere ting til hverandre og for å spisse blyantene sine. Jeg snakket med de fleste elevene. Jeg spurte om de syntes det var gøy å gå på skolen og de sa alle ja. Et par av dem sa at de likte seg bedre her fordi de kjedet seg på den skolen de gikk på før. De andre timene jeg observerte var også preget av ro og elevene hadde fått overlatt mye ansvar til seg selv. De reiste seg ofte for å hente ting og gikk på toalettet uten å spørre læreren.

5.2.2 Observasjoner fra en ”student led conference”

Målet med denne konferansen var at elevene kunne vise sine foreldre hva de hadde gjort det siste halve året. Alle klasserommene hadde gjort klart sine program. I noen klasserom hadde elevene forberedt en mappe der de hadde skrevet et brev til sine foreldre om hva de hadde gjort som de var stolte av og hvorfor. Læreren befant seg i klasserommet sammen med elevene som tok med seg foreldrene inn for å vise det de hadde laget og gjort i løpet av det siste halve året. De fleste foreldre deltok. Barnekoret sang for alle foreldrene og det ble servert doughnuts og saft til alle.

5.3 Øvrige observasjoner i skole C, D og E

Jeg vil ikke gå detaljert inn på de andre offentlige skolene, men beskrive noen av de observasjonene som var sammenfallende med skole B. Alle tre skoler var også preget av

”vurdering FOR læring” ved hjelp av tydelighet og kriteriesetting, men brukte ulike metoder til å gjennomføre dette.

Skole C hadde leseforståelse som sitt ”instructional focus” og jobbet mye med dette i samarbeid med skolebiblioteket. I denne skolen gjorde jeg flere klasseromobservasjoner og så en tydelig systematikk i metoder for egenvurdering, vurdering av hverandre og felles kriteriesetting. En annen observasjon var at flere klasser brukte fargene rød, gul og grønn for å vise deres forståelse i faget. Grønn betydde at de forsto, gul betydde at de ikke forsto helt og rød betydde at de ikke forsto i det hele tatt. De brukte disse plastkoppene i timene for å synliggjøre for læreren, uten å rekke opp hånden, om de forsto eller ikke. Dette var også et virkemiddel som ble brukt for at elevene kunne hjelpe hverandre. Dersom en elev hadde satt opp den røde koppen, kom elever med grønn kopp bort for å hjelpe dem. Lærerne var veldig opptatt av å knytte læreplanmålene til det som ble gjort i timene og kriterier og læringsmål var hengt opp på vegger i ganger og klasserom. De brukte ”rubrics” i stor grad, som var vurderingsskjemaer der eleven selv eller læreren ”sjekket” om vurderingskriterier var nådd

I skole D så jeg også tydelig ”vurdering FOR læring” i praksis. Jeg observerte både det vanlige skoleprogrammet og ”Child study program” som er et program som er knyttet til forskningsprosjekter ved Universitetet i Alberta. I det vanlige skoleprogrammet observerte jeg en 5. klasse der elevene skulle fremføre små historier. De satt i ring og lyttet til en og en som leste opp sin historie. Elevene skrev notater og spørsmål på små kort som de kalte ”Q-cards” som de brukte i sin respons til eleven som framførte sin historie. Deres respons er knyttet til kriterier som de selv har satt i felleskap i timen før. Det er stor grad av disiplin. Elevene irrettesetter hverandre dersom det er småsnakk og uroligheter.

Skole E som hadde ”vekst i forståelse” som ”instructional focus”, var også opptatt av ”vurdering FOR læring” i stor grad. De brukte mye ”jeg kan... ” setninger der de sammen gikk igjennom hva de hadde lært og hva de kunne. Det var stort fokus på kriteriesetting og definering av læringsmål og alt var tydeliggjort på veggene også her. I følge Black og Wiliam, har formative vurderingsstrategier spesiell god innvirkning på de med lavest nivå (jf Black og Wiliam 1998). Denne skolen hadde ca 45 % elever med spesielle vansker og brukte de samme metodene for disse elevene.

Alt i alt tyder observasjonene på at alle de fire offentlige skolene som var drevet av EPSB (skole B, C, D og E) hadde stor innsikt i ”vurdering FOR læring”. I skole A var fokuset mer på ”vurdering AV læring”.

6 Analyse

I denne delen vil jeg besvare problemstilling nummer en: *Hvordan preger Albertas vurderingskultur læreres bruk av vurdering som profesjonsfaglig kompetanse og som verktøy for læring i to barneskoler i Edmonton?* Jeg definerer begrepet profesjonsfaglig kompetanse ved å ta utgangspunkt i Dales fire kunnskapselementer som utgjør lærerens profesjonsgrunnlag (Dale 1999). Avslutningsvis gjør jeg en redegjørelse for hvordan skole A og B bruker vurdering som verktøy for læring før jeg tar for meg ”skolekoden” i disse skolene.

6.1 Vurdering som profesjonsfaglig kompetanse

Jeg vil her bruke intervjuer fra skole A og B til å belyse hvordan lærerne på disse to skolene bruker vurdering som profesjonsfaglig kompetanse. Jeg tar utgangspunkt i disse fire spørsmålene: 1) *Hva slags kunnskap lærerne har om vurdering?* 2) *Hvordan spesifiseres kriteriene, målene og innholdet for vurderingen i skolen?* 3) *Finnes det noen felles retningslinjer for vurdering i lærerkollegiet?* og 4) *Hvordan utvikles kriterier for vurdering som skal fremme kvalitet i deres skole?* Jeg henviser til intervjuvarene direkte på engelsk for å få en mest mulig nøyaktig beskrivelse av hvordan lærerne tenker om vurdering.

6.1.1 Hva slags kunnskap har lærerne om vurdering?

Alle de tre lærerne som ble intervjuet på skole A har tatt kurs om vurdering i forbindelse med lærerutdanningen ved universitetet i Alberta. De har også tatt diverse ”workshops” som omhandler vurderingsstrategier utenom utdanningen. De to kvinnelige lærerne (A1 og A2) svarte at de synes de lærer best av andre lærere og gjennom erfaring, mens den mannlige læreren (A3) følte at det kunne vært interessant med mer kunnskap om vurdering. Han var nå i gang med en bok av en amerikansk forfatter som han også hadde deltatt på kurs med. - *Right now I am looking at a kind of assessment that assesses kids in terms of what they call neural development functions.* Han syntes dette var spennende og kunne tenke seg å lære mer om lignende strategier. De to kvinnelige lærerne følte ikke spesielt behov for flere kurs og workshops, noe den eldste kvinnen (A2) ga uttrykk for ved å si: - *For me I do not think that more education is valuable.* Den yngre kvinnelige læreren (A1) sa: - *because at this point I am kind of specialized into science and my main focus is teaching science and a lot of*

the assessment is very language arts based and very individual group learning, and our focus is "whole group learning". Med andre ord passet ikke innholdet i de kursene hun hadde vært på til de metoder hun brukte i sin klasse.

Alle tre lærerne i skole A hadde kunnskap om at en overfokusering på "vurdering AV læring" kan ha negative konsekvenser, men overskygget dette med å trekke frem positive sider med resultatvurdering. Den mannlige læreren (A3) fortalte om en rektor han kjente i Los Angeles som sammenlignet formålet med vurdering med gjøremål i militæret. Gjøremålene var bare noe som måtte gjennomføres selv om man egentlig ikke hadde lyst til det. "Level of concern" var det viktige poenget her, det vil si å få opp konsentrasjonen. Han sa: - *Fear can help you to prepare. If you're afraid of being evaluated by your peers in a speech well that can motivate you to do better and to study.* Likevel var han klar over at frykt kunne ødelegge for læring og visste om forskning som viste til dette, men at det gjaldt mer innlæring av "dype" ting. - *You can be worried and that is ok with the superficial kinds of things that you need to learn, but the more complex kinds of skills, art and other knowledge and things like that, should come from inquiry that is free from fear. But that is the deeper stuff, you can still get to the other stuff do not worry.* Den yngste kvinnelige læreren (A1) sa at den mest positive siden med vurdering var: - *It keeps them on their toes.* Hun fortsetter: - *I mean we live in a world unfortunately where everyone wants something from something else, so by getting a mark or some kind of feedback, it continually helps them to improve.* Den eldste kvinnelige læreren (A2) sa at fordelene med vurdering var at hun vurderte seg selv som lærer og at alle lærere trenger det. - *If a child do poorly then you need to look at what you have done.* Med andre ord så hun resultatene som en vurdering av seg selv og sin undervisning.

De andre lærerne jeg snakket med utover mine intervjuinformanter i skole A, samlet også informasjon om elevens kunnskapsnivå ved hjelp av resultatbaserte prøver. Dette preget deres undervisningsmetoder og deres pedagogiske syn. Visjonen for skolen som var tydeliggjort på deres nettside og i skolens håndbok for foreldrene viste at de er opptatt av en tradisjonell lærersentrert undervisning der disiplin og autoritet spiller en viktig rolle. Deres visjon har de formulert slik: "Provide an orderly and structured environment, with a properly

sequenced teacher - directed instruction and strong home / school partnerships, where average children can excel in an academically orientated program”²⁵.

Som nevnt tidligere var skole A rangert som nr 30 blant alle skoler i Alberta i en sammenligning av resultatene i de provinsielle prøvene. Det virket som om oversikten over skolenes rangering fungerte som en slags motivasjon for både elever og lærere fordi det ble sett på som et bevis på deres akademiske suksess i provinsen. Den eldste kvinnelige læreren (A2) på skole A sier i intervjuet at: *-The formal evaluation done by the province is something that is in the back of our heads all the time. Particularly those who teach grade 3, 6, 9 and 12, because we know our school placement in Alberta.*

Da jeg spurte rektoren og lærerne på skole B hvilken plass de var rangert til, visste de det ikke. De var heller ikke opptatt av slike rangeringer fordi de mente at resultatene i de provinsielle prøvene ikke kunne være et mål på hvor god en skole var. Jeg ble fortalt av rektoren på skole B at det alltid var de private skolene og charterskolene som lå på topp i disse rankinglistene fordi mange av dem ikke tok inn elever med spesielle vansker. I skole B så verken lærerne eller rektoren en slik rangering som et mål på suksess. I stedet var de opptatt av at skolen utviklet seg de siste årene i en positiv retning og at de hadde forbedret elevenes resultater, spesielt i leseforståelse.

De tre lærerne jeg intervjuet på skole B, hadde også tatt sin utdannelse ved Universitetet i Alberta og hadde i tillegg tatt workshops som omhandlet vurderingsstrategier. Til tross for samme utdannelse jobbet de annerledes med vurdering i skole A enn i skole B. Skole B var også opptatt av gode resultater på deres skole, men deres middel til å nå dette var ikke en overfokusering på resultatbaserte prøver. I tillegg til de provinsielle prøvene brukte lærerne andre vurderingsformer underveis i læringsprosessen fordi de mente at resultater ikke kan gi et riktig bilde av en elevs læringsutbytte. De var alle tre opptatte av de negative konsekvensene de provinsielle prøvene kunne føre til og pekte på at frykt kunne føre til lavere motivasjon hos elevene. Da jeg spurte dem om det var noen positive sider med slike prøver, var de mer opptatt av ulempene med slike tester. Den nyutdannede mannlige læreren (B3) sa - *I think it is a disadvantage to use it to grade teachers, to evaluate teachers and rank school and to answer parents.* En positiv side med vurdering var at det kunne brukes som et ståsted for læreren og eleven. Den ene kvinnelige læreren (B2) som underviste i det

²⁵ Fra “Introduction to ... Charter School for prospective parents” (2006)

alminnelige programmet sa: - *I do not want the students to be fearful of the tests, I want them to welcome the tests. - The test could always be an opportunity to prove to yourself first, how far you have come.* Den kvinnelige læreren (B1) som underviste elever i programmet for elever som lå høyt over gjennomsnittet, var veldig opptatt av at andre metoder skulle brukes. Hun sa: - *Is assessment just the test? Because I do not believe in tests. Tests are only a little tiny capsule of what the child does. And if you have tests and anxiety then you do very poorly on a test. It is not a demonstration of what you learn or what you know.*

Som en oppsummering vil jeg konkludere med at lærerne på skole A og B hadde god kunnskap om vurdering, men har selv valgt strategier og metoder som passer best til deres skole og skolens visjon. De har alle seks samme bakgrunn fra Universitet i Alberta og av den grunn tilgang på kurs om vurderingsmetoder til bruk både underveis i læringsprosessen og ved læringsprosessens slutt.

6.1.2 Hvordan spesifiseres kriterier, målene og innhold for vurderingen?

Jeg vil nå se på hvordan de to skolene forklarer hvem som setter kriterier, mål og innhold for vurderingen. Som nevnt tidligere styres både skole A og skole B av læreplanen og de provinsielle prøvene som styres Alberta Education. Til tross for dette, hadde lærerne på skole A og B ganske ulike oppfatninger av hvordan vurderingskriterier settes.

Da jeg spurte lærerne på skole A om hvem som satte kriteriene for vurderingen i klasserommet, fikk jeg raske tydelige svar. Den eldste læreren (A2) forklarte dette ved å si: - *The curriculum dictates what that child needs to know when I see him or her the first day at school.* Jeg spurte også lærerne om elevens rolle i forhold til å sette kriterier og de ga alle tre uttrykk for at eleven hadde en liten innflytelse i kriteriesettingen. Den mannlige læreren (A3) sa at elevene hadde liten innflytelse, men at han forsøkte å tydeliggjøre læringsinnholdet for dem. Han forklarte dette ved å si: - *I try to get it clear to everybody what it is we are shooting at. (...) Sometimes I have them to evaluate one another. They take a role as an evaluator. But in my school we would take a very cautious approach, and we would do it very methodically and we will control what they are evaluating.* Videre sa han at elevene vurderer seg selv i liten grad. Han tilføyde at var klar over at det var mye av dette i de andre vanlige skolene. Selv forsøkte han å øke elevenes forståelse av hvor deres ståsted er ved å vise dem eksempler på et godt arbeid eller ved å vise hvor langt de andre elevene har

kommet. Han forklarte dette nærmere ved å si: *-I try to encourage them to look with the end in mind, so if we have an assignment of some sort, I would like them to build a picture of what it is supposed to look like when it is finished. And then go ahead from there that helps for self evaluation and organization.* Den yngste kvinnelige læreren (A1) var enig i at det var læreren selv som satte kriteriene ut fra deres egen læreplan: *- For the most part I kind of lay out what I am expecting to be attained, but however there is freedom within that.* Hun tok for seg et eksempel der hun fortalte om en plakat elevene skulle lage i naturfag. Hun lagde selv retningslinjer for hva elevene måtte gjøre for å få over 80 % og hva de måtte gjøre for å få over 70 %, men utenom det kunne de selv ta vekk eller forandre detaljer innenfor rammen av basiskriteriene. Hun avsluttet med å si: *- Somehow they have to complete exactly as I have asked, but there is freedom within that.* Hun fortalte at det hendte at elevene en gang i blant gjorde oppgaver uten at de fikk karakterer. Målet med dette var i følge læreren at elevene kunne få kjenne på hvordan det var å gjøre noe uten å bli målt.

Som jeg har nevnt tidligere var lærerne på skole B veldig opptatt av å knytte det elevene gjorde til læreplanmål. Til tross for at læreplanmålene ble tydeliggjort i stor grad og dette var nært knyttet til skolens vurderingskriterier, var det tydelig større frihet for elevene i vurderingsprosessen i skole B enn i skole A. Når jeg spurte hvem som satte kriteriene svarte den kvinnelige læreren (B2) for det alminnelige programmet: *- We all do. We are all part of the criteria. We do it together.* Videre forklarte hun: *- I want them to be independent and critical thinkers. I want them to be analytical. It is really easy for me to say this is what you need to write on your exam. Then they are writing to satisfy my needs. But I want them to learn to take a stand and satisfy and be independent thinkers. So if they write something for me and that is the fastest way to get an A. It wont work. We have to study things together. And then they have to work it out on their own.* Den andre læreren (B1) for det andre programmet var enig i en slik tenkning og svarte: *We do it together. Sometimes I do it, sometimes they totally do and I sometimes let them assess each others' work. And they really like that idea. And they learn the language. When the grade five comes to me and says I like that because it looks nice. Then I say, well I am sorry, "nice" means nothing. So then they learn. Is it well written? Was it explained? Did you learn something? Were they able to answer the questions? And so on. And they get really into that. (...) So I really believe that the kids should be involved then they have ownership of it so they work harder. Well not necessarily harder, but smarter. And we have smartiegoals, goals that can be measurable, goals that can be achievable and goals that are realistic.* Den nyutdannede læreren (B3) var

også opptatt av å inkludere elevene i å sette kriterier, men hadde ikke fått gjort det så mye, siden dette var hans første år. Han følte seg ikke helt trygg til å overlate ansvaret til elevene, men det var noe han håpet å få gjort mer den kommende høsten. Han sa: *I like the idea of having them to create their own rubric because it is their own and then it is really accountable for them. But it is not something I have been engaged in yet this year.*

Jeg tolker intervjuintervjuene dit hen at lærerne på skole B var opptatt av å inkludere elevene i vurderingsprosessen i større grad enn skole A. Lærere i skole A brukte begrepet ”evaluation” i større grad enn skole B som brukte ”assessment” som et samlebegrep for all vurdering av eleven. Dette ga et inntrykk av at skole A målte verdien av elevenes resultater ved at disse ble målt i prosenter og satt opp mot skolens og provinsens standarder. Skole B var også opptatt av resultater, men brukte flere verktøy underveis i læringsprosessen enn bare karakterer for å styrke elevenes læring. De inkluderte elevene i stor grad, elevene vurderte seg selv og lærerne var opptatt av formative vurderingsstrategier. Dette er i henhold til Black og Wiliam (1998b) som vektlegger egenvurdering og det å inkludere elevene i læringsprosessen ved hjelp av tydelig målsetting for å oppnå effektiv undervisning. I skole A kan vurderingspraksisen tyde på at ineffektiv undervisning kunne forekomme fordi elevene ikke ble inkludert i vurderingsprosessen i stor grad. Deres egen læreplan og vurderingsmetoder ga lite rom for å praktisere de formative vurderingsmetoder som i følge Black og Wiliam (1998a) bidrar til læring (jf del 2.3).

6.1.3 Finnes det noen felles retningslinjer for vurdering i lærerkollegiet?

I skole A hadde lærerne i følge lærer A3, laget sitt eget vurderingssystem som de fleste lærerne fulgte. I engelsk for eksempel deles faget inn i syv underkategorier som alle elevene ble vurdert i. En slik oppdeling satt et preg på deres vurderingsstrategier og de vurderte elevene hver uke i alle fag ved hjelp av prøver og oppgaver. Den yngste kvinnelige læreren (A1) hadde i sin klasse delt opp naturfagskarakteren i tre deler der oppgavene telte 50 %, prøvene 40 % og deltagelse 10 %. Denne oppdelingen kunne variere fra lærer til lærer. Hun sa at hun ofte diskuterte med en annen naturfagslærer hvor mange prosent et prosjekt skulle telle ut fra den endelige karakteren. Den mannlige læreren (A3) mente at de var enige om metodene for ”vurdering AV læring”, som var innsamling av resultater, men de snakket ikke ofte om selve elevvurderingen i klasserommet. Videre tilføyde han at de ikke hadde noen organiserte eller systematiserte samtaler om vurderingsstrategier og metoder felles i

kollegiet. Dette var den yngste kvinnelige (A1) læreren enig i, mens den eldste (A2) sa at de snakket om det ofte, men henviste da til at de diskuterte hvordan de skulle øke resultatene til elevene. Hun nevnte også at de kunne få hjelp av eksperter dersom elevene ikke fikk gode resultater: - *We also have available educational psychologists. So if we have a student that is not successful for whom we don't know the reason, we can refer that child to an educational psychologist who will perform basic intelligence tests dependent on the anecdotal report we give. And he or she may give that child other kinds of tests.* Alt i alt ga lærerne i skole A uttrykk for at de ikke snakket mye om andre vurderingsstrategier enn de resultatbaserte vurderingene og hvordan de kunne få elevene til å prestere bedre i disse testene. De brukte alle "reportcards" der de forklarte resultater i prosentandeler. Den yngste læreren (A3) tok frem et "report card" til en flink elev og forklarte bruken av det: - *I mean... this student has 92, 96, 100 and 98 %. If she wants to work anywhere it is obvious that she can do a little bit more work in her assignments, but for some students it is evident, here again it is participation. Its bringing things in, being prepared and having what you need to be ready. So it kind off breaks down where we need to focus our energy for improvement.*

Da jeg spurte de tre lærerne på skole B om de visste hvilke vurderingsstrategier de andre lærerne på skolen brukte, svarte den nyutdannede læreren (B3): - *Well I have a mentor teacher that I work with and I got some rubrics from her. But I have not really talked to any teachers about how they assess in their classroom.* De to kvinnelige lærerne var klar over hvordan de andre jobbet med vurdering i klasserommet. Læreren (B2) i det vanlige programmet svarte: - *Well we always get together. We discuss what is happening from one grade to another. There is very much a collaborate project at this school (...) I don't think there is too much difference in how we assess. There might be some differences in whether we are in the gifted or regular program. However the goals as a school are the same.* Den andre læreren (B1) for det andre programmet, hadde også kunnskap om hvordan de andre jobbet med vurdering i klasserommet og sa at det varierte ut fra ulike klasser og situasjoner. Hun mente at det var en del ulike strategier, men at de ofte snakket om dem i fellesskap. Jeg ble også presentert for en bok om vurdering som mange av lærerne jobbet med.

En dag i måneden møttes alle lærerne i skole B for å diskutere skolens strategier og metoder for å oppnå profesjonell utvikling. Jeg deltok på et slikt møte der de gikk igjennom punkter som skulle føre til "professional development". Disse punktene var: 1) bygge ekspertise, 2) sikre endring i praksis, 3) kommunisere effektivt og 4) styrke elevenes læring. Sammen gikk

de igjennom punktene for å finne “bevis” på at de alle jobbet med disse punktene. Under punkt nummer 4 som gikk på å styrke elevenes læring, jobbet de med følgende punkter: kollegial analyse av resultater i distriktstesten, rubrikker som er relatert til et spesifikt fagområde og som blir brukt og redefinert kontinuerlig, læreplanens forventninger uttrykkes tydelig og elevenes arbeid støtter disse forventningene og de har en pågående “assessment for learning”²⁶. Alle disse punktene observerte jeg selv på skolen og det var interessant å se at den teorien de ønsket å bruke faktisk ble gjennomført i praksis. Etter møtet med lærerne møtte rektoren for skole B rektorer fra andre skoler i distriktet. Slik møter hadde de ca en gang i måneden. I disse rektormøtene som de kalte “principal support team”, delte rektorene erfaringer fra sine skoler og knyttet det til ”professional development communities” som jeg beskrev i del 4.3. Referatet fra dette møtet ble videreført til EPSB slik at de kunne kommunisere sine behov til ledelsen i EPSB. Skole C, D og E jobbet også aktivt med ”vurdering FOR læring” i ”Professional learning communities”.

Begrepet ”Professional learning communities” bygger på Peter Senges (1990) definisjon på lærende organisasjoner: “...organizations where people continually expand their capacity to create the results they truly desire, where new and expansive patterns of thinking are nurtured, where collective aspiration is set free, and where people are continually learning to see the whole together” (Senge 2004). I følge Senge (1994) krever en lærende organisasjon at medlemmene i organisasjonen har tilgang til ressurser som tid til å samarbeide, pågående støtte fra ledelsen, informasjon og klar tilgang til kollegene. Mangel på meningsfulle muligheter til å delta i læringsaktiviteter kan begrense muligheten for en skole å bli en lærende organisasjon (Lashway, 1997 i Roberts and Pruitt, 2003).

Selv om Senges teori har vært utsatt for kritikk²⁷ fordi det er vanskelig å overføre denne teorien om lærende organisasjoner i praksis, ble den i flere skoler i Edmonton brukt som en tankemodell for hvordan man utvikler ”professional learning communities”. To innvendinger til Senges teori er at det er vanskelig å finne organisasjoner som har de karakteristikkene som Senge mener identifiserer en lærende organisasjon. I vårt kapitalistiske samfunn vil en visjon om at organisasjoner og bedrifter skal helhjertet kultivere læring bli begrenset (Klein 2001). Videre er ikke hans teori satt inn i noen politisk og moral kontekst og hans tilnærming opererer bare på bakgrunn av organisatoriske interesser (Klein 2001). Til tross for disse

²⁶ Tatt fra programmet for et lærermøte jeg deltok på

²⁷ Se: www.infed.org/thinkers/senge/htm

innvendingene brukte skolene i EPSB "vurdering FOR læring" som strategi for å nå sin felles visjon som var læring for elevene. De jobbet sammen om dette i kollegiet innad i skolene, på tvers av skolene og på tvers av nivåene i utdanningssystemet. Det var et tett samarbeid mellom blant annet "Alberta Assessment Consortium" (AAC), som er en uavhengig organisasjon som lager vurderingsverktøy for skoler og universiteter, "Alberta Teacher's Association" (ATA) og Universitetet i Alberta som alle var opptatt av "vurdering FOR læring".

Både skole A og B hadde felles retningslinjer for "vurdering AV læring" blant sine kolleger. Dette skyldtes først og fremst de obligatoriske provinsielle prøvene. Skole A hadde også ukentlig resultatbaserte prøver utover de provinsielle prøvene. Intervjuene og observasjonene tydet på at et slikt fokus på "vurdering AV læring" preget lærerens oppfatning av vurdering. De hadde ikke satt "vurdering FOR læring" i system eller hadde heller ingen felles retningslinjer for dette. "Vurdering AV læring" derimot hadde de systematisert i stor grad ved å knytte vurderingen til mål i sin egen læreplan. I skole B fikk lærerne frigjort tid til møter der de diskuterte hvordan de sammen kunne delta i et profesjonelt læringsmiljø. For å oppnå dette var blant annet "vurdering FOR læring" og bruk av "rubrics" satt som to mål som de skulle jobbe med i fellesskap. Skole A hadde bare felles retningslinjer for "vurdering AV læring", mens skole B hadde felles retningslinjer for både vurdering AV og FOR læring.

6.1.4 Hvordan utvikles kriterier for vurdering som skal fremme kvalitet i skolen?

I skole A tydet erfaringene på at gode resultater er et kriterium for å fremme kvalitet i deres skole. Målet er å vise til gode resultater og en god rangering i forhold til de andre private skolene og charterskolene som ligger resultatmessig på topp i Alberta. Et slikt fokus har ført til en tradisjonell lærerstyrt undervisning der læreren er autoritær og instruerende. Mine observasjoner og intervjuer tyder på at lærerne på denne skolen var preget av en "teaching to the test" mentalitet. Elevene blir lite inkludert i vurderingsprosessen og de lærer det de vet de må lære for å bestå prøvene. Foreldrene ønsket gode resultater på prøvene og det var det lærerne jobbet for å skaffe (jf samtale om guttens resultater i "foreldre – lærer – intervju" s. 49). Som den ene unge kvinnelige læreren (A1) forklarte, lever vi i et samfunn der alle forventer noe av hverandre. Foreldre sender sine barn til denne skolen og forventer gode resultater, og lærerne jobber hardt med å tilfredstille foreldrene ved å vise til elevenes sluttresultater.

I skole A var ”vurdering AV læring” blitt et formål i seg selv der resultatene ble en motivasjonsfaktor for både elever og lærere. Den eldste kvinnelige læreren (A2) på skole A forklarer dette med å si: - *If I am going to be an adequate teacher I need to determine whether they are prepared to carry on with it. Have they reached the level or not. You need to have these assessments, otherwise you will be talking to the wall and nobody will understand you.* For å nå høye resultater med ”whole group learning” for barn over gjennomsnittet, har skolen en tradisjonell undervisningsform som er målstyrt i stor grad. De baserer seg på en mål - middel pedagogikk der et stort fokus på ”vurdering AV læring” er et middel til å nå målet, som er gode resultater.

I skole B var lærerne også opptatt av gode resultater spesielt i leseforståelse fordi dette var deres ”instructional focus”. Likevel tyder intervjuintervjuene og observasjonene på at deres kriterier for å fremme kvalitet var noe annerledes enn i skole A. EPSB har satt en del kriterier til sine skoler som de må følge. Disse kriteriene innebar metoder for vurdering FOR og AV læring og metoder for å øke den profesjonelle kompetansen (jf ”professional development communities”). Mitt inntrykk etter å ha besøkt skole B, C, D og E, som alle var en del av EPSB, var at de jobbet med å øke sin profesjonelle kompetanse på individ -, klasse - og systemnivå, for igjen bedre kvaliteten. Skolen ønsket gode resultater, men jobbet med læringsprosessene hos både elever og lærere for å få til dette.

EPSB har lagt føringer for hvordan skole B, C, D og E jobber ved å tilrettelegge for alternative metoder for både vurdering AV og FOR læring. Kreativitet og nytenkning er verdsatt i stor grad og skolene jobber i nettverk for å oppnå lærende skolemiljøer slik at lærere i større grad kan bruke sin profesjonelle kompetanse for å imøtekomme elevenes behov. Både skole A og B er målorienterte, men middelet til å nå målet om gode resultater er ulikt. Skole B jobbet for å inkludere elevene i læringsprosessen slik at de kunne jobbe ”smartere”, mens skole A brukte resultater underveis til å motivere elevene til å jobbe ”hardere”.

6.2 Vurdering som verktøy for læring i klasserommet

Jeg vil nå se hvordan begge skolene jobbet med ”vurdering FOR læring” i klasserommet gjennom egenvurdering, vurdering av elevene imellom og involvering av elevene i kriteriesetting.

Begge skoler brukte vurdering som et verktøy, men på ulike måter og ut fra ulike formål. Skole A brukte vurdering som et verktøy til å informere elevene, foreldrene og seg selv om hvor bra elevene gjør det. Som den mannlige læreren (A3) sa: (...) - *I am required to give a kind of order that shows who is on the bottom and who is on the top. So that they can get a percentage on the report cards, based on what they have done.* I skole B ble vurdering brukt til å informere elevene, lærere og foresatte, men de brukte i tillegg andre metoder og strategier for å samle informasjon om eleven. Lærerne i skole B forklarte at det var nødvendig å ta hensyn til de provinsielle prøvene fordi det er flere parter i provinsen som ønsker informasjon om resultatene. Den kvinnelige læreren (B2) for det vanlige programmet i skole B forklarte dette dilemmaet slik: - *I guess I have to decide whether I do the assessment for the school or for the child, the assessment for the classroom, or for the province. First of all I think the assessment should be for the student first. Simply to know where you start from, what you need to work with and where you need to go. Then I have assessment goals that I have to be met with within the classroom for the school, for the school district and for the province.* I skole B var aktivitetene som ble brukt for å skaffe informasjon om elevens læring varierte og ikke bare preget av resultatbaserte prøver. I skole A ble informasjonen for det meste samlet inn ved hjelp av resultatbaserte prøver for å imøtekomme foreldrenes behov for synlige resultater og for å kunne skåre høyest mulig på de provinsielle prøvene.

I skole A var ikke ”vurdering FOR læring” et begrep som ble brukt, fordi som alle lærerne selv uttrykket det var det ikke mye av denne formen for vurdering på denne skolen. Lærerne i skole A forklarte det med at det ikke passet til hvordan de jobbet på deres skole. Måten den mannlige læreren (A3) ville få dem til å reflektere rundt sitt eget arbeid og ståsted på, var å vise til andre elevers arbeid. Med andre ord ble konkurranse brukt som en motivasjonsfaktor. Her igjen brukes frykt som en slags motivasjonsfaktor fordi elevene er redde for å stryke eller ikke å gjøre det bra nok. Imidlertid fikk elevene lov til å evaluere hverandre i noen grad og sammenligne sine resultater.

I skole B var ”assessment FOR learning” et velkjent begrep som alle lærerne var inneforstått med. Alle lærerne jeg pratet med ga uttrykk for at resultater ikke ga et helhetlig bilde av hva eleven kan. ”Vurdering FOR læring” var noe alle jobbet for slik at elevene skulle lære bedre. Jeg observerte at elevene vurderte seg selv og hverandre og de satte kriterier i fellesskap. Fargene rød, gul og grønn som ble brukt for å vise sin forståelse i oppgaver de jobbet med.

Lærere kalte prøver for quiz for å komme vekk fra frykten for prøver og man gikk igjennom hva man hadde lært punkt for punkt i løpet av en time. Mye av denne undervisningspraksisen er i henhold til Black og Wiliams forslag til metoder for formativ vurdering (jf del 2.3). I skole C, D og E observerte jeg det samme fokuset på ”vurdering FOR læring”.

6.3 Hva sier intervjuene om ”skolekoden” på skole A og skole B?

Intervjuene og observasjonene tyder på at Albertas vurderingskultur som i flere år har vært preget av ”accountability” og privatisering preger vurderingskulturen på to ulike måter og det kan være flere årsaker til dette. En forklaring på at disse to skolene har ulike oppfatninger av formålet med vurdering vil jeg forklare med begrepene ”skolekode” og rammefaktorer. ”En ”skolekode” består av en samling styrende prinsipper for handling og tolkning; disse omfatter det som er viktig når det gjelder arbeid, arbeidsmiljø og generelle skoleproblemer” (Arfwedson i Engelsen 2006:247). Det betyr at det på hver skole ”sitter noe i veggene” som virker inn på alle lærernes tenkning og arbeid. De utvikler holdninger til og synspunkter på pedagogiske spørsmål, måter å forholde seg til elever på, daglige rutiner osv. som er forskjellig fra dem man finner på andre skoler (Engelsen 2006). ”Skolekoden” er noe tydeligere i skole A enn i skole B. Dette gjenspeiles både i intervjuene og i mine observasjoner. Erfaringene kan tyde på at den nyliberale markedstenkningen fortsatt sitter igjen i veggene i denne skolen. Gode resultater var det de jobbet for og både lærere og elever så på gode resultater som suksess. I skole B så man på framgang som suksess, noe som galdt hver enkelt elev og skolen som felleskap. Skolen var opptatt av å sammenligne seg med seg selv og lærere var opptatt av å lære gode læringsmetoder fra andre lærere slik at de kunne forbedre seg.

En annen viktig årsak til at skolene er så forskjellige er rammene. I følge Lundgren (1979 i Engelsen 2006) er det liten tvil om at rammefaktorene er relevante for læreres arbeid. Lundgren mener at enkelt lærere ikke kan endre skole og undervisning uavhengig av rammer. Undervisning foregår innenfor rammen av samfunnets økonomiske, sosiale og kulturelle struktur, som tildels kan forklare virksomheten i skole og klasserom (Lundgren 1979 i Engelsen 2006). Dette gjør at det ikke alltid er mulig å bruke den didaktiske relasjonsmodellen som intendert i alle undervisningssituasjoner fordi rammene er styrende i

stor grad. Lundgren mener at dersom man ønsker å endre undervisningen, nytter det ikke med henvendelser til enkeltlærere eller med endringer i lærerutdanning og læreplaner, man må i stedet endre rammevilkårene for skole og undervisning. Lundgren rettet primært oppmerksomheten mot de mer formelle rammebetingelsene for skole og opplæring – rammebetingelser som kunne bli ført til sentrale utdanningsmyndigheters vedtak. Lundgren har blitt kritisert for dette blant annet fordi også andre uformelle rammefaktorer kan virke inn på de valg lærere tar (Engelsen 2006: 246). Mine erfaringer fra skole A og B tyder på at Lundgren har rett med tanke på at de formelle rammebetingelsene påvirker læreres undervisningspraksis og handlingsrom. Både skole A og skole B følger samme læreplan og provinsielle vurderingssystem, og dette preger deres vurderingspraksis. Likevel tyder erfaringene på at lærernes og ledelsens egne tolkninger av de mulighetene og de begrensningene rammebetingelsene "Alberta Education" gir har også mye å si for deres undervisningspraksis. Gjennom sin oppfatning av mål og andre betingelser kan lærerne ha sterk innvirkning på forløpet av undervisningsprosessen. Videre vil lærernes pedagogiske og faglige syn også virke inn. Dette er i henhold til Svingby som mener at lærerne innenfor rammefaktorene har et visst handlingsrom. Hvordan lærere velger å utnytte sitt handlingsrom er avhengig av deres pedagogiske grunnsyn og praktiske undervisningsteori (Svingby 1978 i Engelsen 2006:246).

Foreldrestyret til skole A og EPSB som styrer skole B, har innenfor rammene til "Alberta Education" har lagt egne rammer for lærerne på sine skoler. Skole A har brukt sine rammer til sin fordel for å støtte resultatorientering og fokus på effektivitet, noe som har ført til en autoritær og tradisjonell lærersentrert pedagogikk. De lar blant annet de provinsielle testene styre deres handlingsrom og undervisningspraksis. Skole B har bevisst utnyttet sitt handlingsrom ved å finne alternative måter for å bedre undervisningen innenfor de rammer som EPSB har lagt. De ønsker en pedagogikk som ikke er preget av resultatbaserte prøver og har fått hjelp til dette fra politisk hold. EPSB har i samarbeid med ATA, ACC og Universitet i Alberta, bidratt til å gi lærere og rektorer verktøy og hjelpemidler til å fokusere på formativ vurdering. Videre er kreativitet verdsatt i høy grad av EPSB og læringserfaringer deles mellom skolene for å skape profesjonelle læringsmiljøer. Begge skoler har en tydelig filosofi som preger lærernes handlingsrom som igjen preger deres vurderingspraksis og undervisnings strategier. Slik jeg ser det kan både Lundgrens og Svingbys rammefaktorteori og Arfwedsons skolekodebegrep være med på å forklare den ulike vurderingskulturen på skole A og B.

6.4 Oppsummering

På bakgrunn av erfaringene fra skole A og B, vil oppsummere hvordan Albertas vurderingskultur preger læreres bruk av vurdering som profesjonsfaglig kompetanse og verktøy for læring i skole A og B. Jeg vil først oppsummere i hvilken grad lærerne i skole A og B bruker vurdering som profesjonsfaglig kompetanse og som verktøy for læring.

Når det kommer til kunnskapen lærerne i skole A og B har om vurdering, har alle de seks lærerne samme bakgrunnskunnskap fra lærerstudiet fra universitetet i Alberta. Lærere i skole A og B har alle tilgang til vurderingskurs, men skole B har vært mer opptatt av å følge med på kurs som tar for seg metoder for ”vurdering FOR læring”, enn lærere på skole A. Kriterier, innhold og mål for vurderingen i fagene bestemmes ulikt i skole A og B. Skole A har tydeliggjort læringsmålene i alle fag, ved å lage enda tydeligere og presise mål i sin egen læreplan som lettere kan være gjenstand for vurdering. Dette gjorde de ikke i skole B. Læringsmål og kriterier er ikke tydeliggjort for elevene på vegger og tavle i skole A, slik som i skole B. Gjennomgående bruker skole A ”vurdering AV læring” som metode for å vurdere elevens læringsresultater. Skole B tydeliggjør også læreplanmålene i stor grad, men formålet med dette er å inkludere elevene i læringsprosessen ved hjelp av blant annet egenvurdering. Kriterier og læringsmål henger fremme i klasserommene sammen med egenvurderingsskjemaer der elevene vurderer arbeid de selv har gjort. Både skole A og B har felles retningslinjer for vurdering blant lærerkollegiet. Skole B er opptatt av varierte metoder for ”vurdering FOR læring” og utveksler erfaringer med andre skoler og lærere ved hjelp av ”professional learning communities”. Skole A har ikke et slikt systematisk samarbeid med andre skoler og bruker først og fremst resultatbaserte prøver som vurderingsform. Kriterier for vurdering som skal fremme kvalitet i begge skolene bestemmes først og fremst av ”Alberta Education”, så av EPSB i skole B og foreldrestyret til skole A. Videre tolker lærerne og ledelsen på skole A og B disse rammene i retning av skolens formål og ”instructional focus”. Dette kan være en medvirkende årsak til at ”skolekoden” og vurderingskulturen i skole A og B er forskjellig.

Jeg vil her henviser til Rubin og Rubin (1995:20) som følgende om kultur: ”Culture is created and evolves to solve shared problems”. Dette er i tråd med definisjonen av kultur som jeg nevnte innledningsvis. Denne oppfatningen av kultur kan forklare hvorfor vurderingskulturen i skole A og B er så forskjellige. Både skole A og B må hankses med de

provinsielle prøvene i provinsen som er felles for alle skolene, men de gjør dette forskjellig. Måten de bruker prøvene på er med på å skape to ulike vurderingskulturer (jf kulturen i skole A og i skole B). Disse ulike kulturene setter ulike grenser for hvordan vurdering kan brukes som profesjonsfaglig kompetanse og som verktøy for læring.

7 Hvordan kan erfaringene fra Edmonton berike den norske vurderingsdebatten?

Jeg vil i denne delen besvare problemstilling nr. 2: *Hvordan kan erfaringene fra Edmonton berike vurderingsdebatten i Norge?* Jeg vil starte med å redegjøre for noen medvirkende årsaker til at vurderingskulturen i Norge ikke er like etablert som vurderingskulturen i Alberta. Jeg vil her ta utgangspunkt i noen faktorer som har påvirket den norske vurderingskulturen, der man har vegret seg for å vurdere hverandre. I følge Gundem kan vår likhetsideologi, vår enhetsskoletanke og frykt for å sette merkelapper på hverandre som angir verdi, være en årsak til at vurderingsbegrepet har vært i skyggen i lang tid (Gundem 1998). Jeg vil henviser til noen undersøkelser som viser tendenser til en slags ”vurderingsvegring” i mange norske klasserom og i lærerutdanningen.

7.1 Trekk ved vurderingskulturen i Norge

I antropologisk forskning har likhet og likeverd lenge vært en spesielt viktig kode i Norge. Antropologen Lidén (2001) mener at oppfatninger om likhet i Norge styrkes indirekte av ulike former for institusjonsbygging som har likhet som mer eller mindre eksplisitte formål. Den norske enhetsskolen er et eksempel på dette. Dette er i tråd med den sosialdemokratiske velferdsstaten som jeg var inne på innledningsvis i oppgaven (jf del 1). Skolens muligheter skulle være like for alle, ved at man skulle likestille forholdet mellom by og land, majoriteter og minoriteter, ulike sosiale grupper, mellom kjønn og vanlige og funksjonshemmede elever (Telhaug 1997). I over 150 år har tanken om en felles skole for alle vært sterk i Norge og det å oppnå enhet og forbedre fellesskapet i skolen har vært en grunnleggende ideologi som alle skolepolitikere i prinsippet har sluttet seg til (Engelsen 2003).

Likhetstankegangen i Norge har også preget lærerutdanningen i Norge. Dette kommer spesielt til uttrykk i den antiautoritære dialogpedagogikken fra slutten av 1970 tallet. Denne ideologien ble preget av en pedagogikk der dialogen var lærestoff på linje med det som sto i lærebøkene. Innholdet i dialogen skulle knyttes direkte til elevens situasjon, behov og interesser og ble valgt ut fra en felles beslutning. Læreren skulle etterstrebe en tilbakeholdenhet som skapte et frihetsspillerom for elevene og ga dem anledning til selvvirksomhet i samsvar med målestokker for et demokratisk samfunn (Dale og Wærness

2006). Denne tilbaketrukne lærerrollen kan man også se spor av i den norske lærerutdanningen i for eksempel veiledningsbøkene til Handal og Lauvås (1999) som har hatt en viktig posisjon i norsk lærerutdanning. Deres veiledningsteori bygger på Torgeir Bues bok som ble utgitt i 1973. Bue (i Skagen 2000) bygger for en stor del på Carl Rogers' modeller for klientsentrert terapi, der kommunikasjonen er et vesentlig innslag. Før denne boken ble utgitt var veiledning bare omtalt som en del av større didaktiske fremstillinger, men med Handal og Lauvås ble veiledning begrunnet mer som en selvstendig didaktisk aktivitet. Handal og Lauvås er opptatt av at lærer og lærerstudent skal ha et symmetrisk forhold, der lærerstudenten ved hjelp av en dialog med læreren kommer frem til sin egen praksisteori uten å bli instruert av læreren (Skagen 2000). Dette er i tråd med den nordeuropeiske didaktikk tradisjonen som setter den reflekterte og autonome læreren i fokus (jf Gundem og Hopman 2002). Et slikt syn på veiledning i lærerutdanningen kan ha ført til en diffus vurderingssystematikk fordi man har vært redd for å sette standarder for hva god undervisning og hva god vurdering er. Videre kan det ha ført til en reservasjon overfor målrelatert vurdering basert på ulike grader av målbeskrivelser noe som forutsetter at målene operasjonaliseres i målbare størrelser (Dale og Wærness 2006).

7.1.1 Behov for ansvarliggjøring også i Norge

I Norge og resten av Europa førte endringene mot slutten av 1970 tallet til at man ga opp troen på offentlig regulering og utjevning og en høyrebølge slo til (Telhaug 1994). Til tross for denne internasjonale påvirkningen beholdt enhetskolen fortsatt sin status i Norge og de utdanningspolitiske endringene var langt fra så markante som i Alberta (jf del 4.1). Likevel er det noen likhetstrekk mellom den politiske utviklingen i Alberta på 80 - tallet og utviklingen i Norge på slutten av 90 - tallet. Omtrent 8 år etter den femdelte business planen som gikk over tre år ble iverksatt i Alberta (jf 4.2), ble Kvalitetsutvalget oppnevnt av Utdannings- og Forskningsdepartementet (UFD) i Norge. Departementet ga følgende mandat for utredningen: ”Utvalget skal i delutredningen beskrive, analysere og vurdere hovedtrekkene i dagens system for kartlegging og vurdering av kvaliteten i grunnopplæringen, herunder rapportering, både på nasjonalt, regionalt og lokalt nivå. Dette må ses i lys av internasjonale prosjekter og initiativ på dette feltet (bl.a. i regi av OECD) og systemer for kvalitetsvurdering i grunnopplæringen som er utviklet i andre land”²⁸. I delutredningen ”Førsteklasses fra førsteklasse” (NOU 2002:10) ble det påpekt at ”Norge er

²⁸ http://odin.dep.no/kd/norsk/dok/andre_dok/nou/045001-020002/hov002-bn.html

et av få land i Vest - Europa som mangler et nasjonalt system for kvalitetsvurdering i grunnopplæringen". På bakgrunn av dette ble utvalget bedt om å utforme forslag til et nasjonalt vurderingssystem der det første hovedområdet skulle omhandle elevenes læringsutbytte. I dette inngår de nasjonale prøvene som skal være en landsomfattende kartlegging av elevenes ferdigheter i skriving, lesing, matematikk og engelsk. Disse prøvene har i følge Utdanningsdirektoratet to hovedmålsetninger: de skal både være et pedagogisk redskap og hjelpemiddel i læring og opplæring og gi informasjon til elever, lærere, skoleledelse, skoleeier, nasjonale myndigheter og allmennheten som grunnlag for dialog og kvalitetsutvikling²⁹. Med andre ord skal innføringen av nasjonale prøver være et hjelpemiddel i å skape et kunnskapsløft i den norske skole. Resultatene i PISA undersøkelsen 2003 viste at norske elever lå rundt gjennomsnittet i de fleste fag og at Norge har et forbedringspotensiale på flere områder (jf del 1.2). Ifølge Lie med flere (i Kjærnsli 2004) er liten bruk av læringsstrategier kombinert med lav konsentrasjon om skolefag, lite forenlig med et godt læringsutbytte, spesielt i en skole der det legges opp til at elevene skal ha mye frihet og ansvar for egen læring. Dersom lærerne ikke er tydelige nok på hvilke læringsstrategier elevene mesterer godt og mindre godt, kan det bli vanskelig for elever å ta del i sin egen læringsprosess. Formålet med "vurdering FOR læring" er nettopp at vurderingskriterier skal bidra til å inkludere elevene i sin egen læringsprosess. Flere undersøkelser tyder på at vi har et forbedringspotensiale på dette området og jeg vil her henwise til disse.

7.1.2 Vurdering som verktøy for læring i norske klasserom

For å få et større innblikk i hvordan vurderingskulturen er i norske klasserom er flere undersøkelser interessante. Et eksempel er rapporten "Lærenes klasseromsarbeid; Interaksjons- og arbeidsformer i norske klasserom etter Reform 97" (Klette, 2003). I følge Klette (2003) er barnetrinnet, især småskoletrinnet, preget av lite irttesettinger. Når det gjelder faglige krav er "kombinasjonen av hyppig bruk av allmenn ros og fravær av eksplisitte og klare faglige standarder for elevenes arbeid, et lite fremtredende trekk i deres klasseromobservasjoner i de skolene som ble undersøkt. Disse funnene kan ikke brukes til generalisering, men er et eksempel på at et utvalg av norske lærere gir få eksplisitte produktkrav og standarder for hva som kjennetegner gode arbeidsbesvarelser og

²⁹ www.utdanningsdirektoratet.no/templates/udir/TM_Artikkel.aspx?id=1400

framføringer” (Klette 2003:52). Differensieringsprosjektet til Dale og Wærness underbygger dette ved at bare 8 % av elevene som svarte i deres prosjekt var helt enige i at lærere ofte forteller dem hva de må arbeide mer med i fagene (Dale og Wærness 2003). I følge Dale og Wærness kan resultatene av Reform 97 sees i lys av en læreplantenkning som tar utgangspunkt i at vi ikke kan sette opp mål som er minstekrav i en skole for alle. Et eksempel på det er Mønsterplanen av 1987 der det sies at elevene må kjenne seg frie til å søke råd og veiledning når de selv føler behov for det. Dette er i følge Dale og Wærness en læreplantenkning som gjør det mulig for læreren å være støttende, inkluderende og positiv, samtidig som de vegrer seg for å stille faglige og eksplisitte krav (Dale og Wærness 2006).

Elevinspektørene 2005 er en annen undersøkelse som viser tendenser som går i retning av at vurdering ikke er tydelig for norske elever. Elevinspektørene som skiftet navn i januar 2006 til Elevundersøkelsen, er en del av det nasjonale kvalitetsvurderingssystemet og ble gjennomført for første gang i 2003. Elevundersøkelsen er en nettbasert spørreundersøkelse der elever kan gi sin vurdering av skolen. Fra og med 2004 ble Elevundersøkelsen gjort obligatorisk for alle elever på 7. 10. og 11. årstrinn. Elevene svarer på spørsmål om motivasjon, trivsel, mobbing, elevmedvirkning, elevdemokrati og fysisk miljø. I analysen av Elevinspektørene 2005 viser Oxford Research viktige funn når det gjelder vurdering i den norske skolen. Når det kommer til elevmedvirkning har 80 % av elevene svart at de medvirker ikke, eller bare i noen fag. Når det gjelder vurdering av eget arbeid har 77 % svart at de medvirker ikke, eller bare i noen fag. Ca 60 % av elevene er uenige i at de er med på å sette egne læringsmål og 1 / 3 sier de er uenige i at undervisningen er tilpasset deres nivå³⁰. Jeg vil her tilføye at spørsmålene som omhandler vurdering i denne undersøkelsen er få og i noen grad upresise. Dette sier også litt om hvordan vurdering som begrep blir prioritert. Likevel kan svarene gi oss innblikk i elevers forståelse og forhold til vurdering, elevmedvirkning og læringsmål.

Prosjektet ”Vurdering som profesjonsfaglig kompetanse”, som jeg har henvist til flere steder i oppgaven, sier noe om et utvalg norske læreres vurderingspraksis. Utgangspunktet for dette arbeidet er nye forventninger og krav til skolen om å utvikle en praksis der vurderingen er et dynamisk og aktivt element i elevenes læringsprosesser (Haugstveit 2005). Utvalget som ble observert og intervjuet i dette prosjektet besto av ni skoler med en lærer pr. skole.

³⁰ http://www.utdanningsdirektoratet.no/upload/Rapporter/elevinspektorene/Som_elevene_ser_det_revidert_Elevinspektorene.pdf

Intervjuszvarene viser blant annet at lærerne savner kunnskap om vurderingsmetoder. Videre viser intervjuene at læreren ønsker analysekompetanse som kan sette skolen i stand til å tolke resultater for å finne ut hvilke tiltak som bør settes inn fra individnivå til institusjonsnivå (Haugstveit 2005). Funn fra min pilotstudie i to barneskoler i Oslo var sammenfallende med funnene i dette prosjektet. Jeg intervjuet tre lærere i to skoler i Oslo (F og G) som alle var opptatt av vurdering og var enige i at formålet er læring og utvikling. Når det kom til kunnskap om vurdering og vurderingsmetoder ga alle de seks lærerne uttrykk for at de ønsket å lære mer om praktisk vurderingsteori. Fem av seks sa at de hadde hatt få eller ingen kurs om vurdering verken i lærerutdanningen eller utenom. Da jeg spurte den ene mannlige læreren (F3) om han følte at han hadde behov for mer kunnskap om strategier for hvordan vurdere elevene, svarte han: - *Ja, det gjør jeg. Jeg har ikke fått det noen steder, det var faktisk i min forrige jobb der jeg var arbeidsleder og hadde ansvar for flere arbeidere, at jeg fikk litt kunnskap om dette. Så den kunnskapen jeg har, har jeg egentlig fått mer gjennom det. Men i lærerutdanningen? Lærte du noe om vurderingsmetoder der? - Nei, ingen ting. Har du hatt noen kurs i vurdering? - Nei.. eller mappevurdering har jeg hatt.. et kurs i forhold til Kunnskapsløftet.*

Alle seks sa at de hadde hatt om vurdering i lærerutdanningen, men at det var i forbindelse med didaktisk teori. De ønsket alle mer kunnskap om vurderingsmetodikk, noe også lærerne i prosjektet "Vurdering som profesjonsfaglig kompetanse" formidlet som et behov (jf Haugstveit 2005). Her er et eksempel fra den andre mannlige læreren (F2) på skole F: Føler du at du har behov for mer kunnskap om strategier og retningslinjer for hvordan vurdere elevene? – *Ja, og ikke teorier, men konkrete tips eller metodikk. Har du tatt noen spesifikke kurs om elevvurdering? - Nei ikke som jeg vet egentlig.. eller elevvurdering har jo vært med i en del fag. Når jeg gikk på lærerhøyskolen så var jo det didaktiske perspektivet med i alle fag, men sånn mer konkret, hvordan skal vi vurdere ... nei egentlig ikke..*

En kvinnelig lærer på skole G (G1) svarte dette da jeg spurte henne hva hun hadde lært om vurdering i lærerutdanningen: - *Da jeg gikk på lærerhøyskolen var didaktikken lagt inn i hvert fag. Men da jeg tok pedagogikk grunnfag så lærte jeg en del om det, men da var det mer overordnet. Vi lærte om hva skjer med eleven, når eleven blir vurdert. Men hva med det praktiske? - Nei ingen ting. Den andre kvinnelige læreren (G2) svarte: - stort sett hver dag har jeg hatt en dialog med alle. Det synes jeg er veldig mye av den daglige vurderingen for da får de kontinuerlig tilbakemelding, men jeg kan veldig lite om vurdering sånn generelt.*

Hvor har du lært de strategier du bruker? - *Jeg vet ikke. Jeg har kanskje lært de på lærerhøyskolen. Eller vi har jo hatt noen kurs som gikk på det med elevsamtalen, og da gikk det på det å se hver elev. Jeg synes det er kjempeviktig at alle føler at de blir sett hver dag. Men jeg har ikke gått på noe rent kurs om elevvurdering. Og hva lærte du om elevvurdering i lærerutdanningen? - Jeg vet at vi hadde noe om det, men det tror jeg ikke var sånn veldig prioritert.*

Intervjuene viste at fem av de seks lærerne hadde lært vurderingsmetoder ved hjelp av erfaring og i samtale med andre kollegaer fremfor i lærerutdanningen. Den kvinnelige læreren (F1) på skole F uttrykte dette med å si: *Det går jo mye på erfaring. Og så litt i utgangspunktet på det at du føler at du klarer å undervise. Det blir jo en sånn blanding at du må jo føle at du når fram og at eleven skjønner hva du mener. Etter hvert så har man jo hatt ganske mange elever... så da vet du sånn omtrent hva eleven kommer til å gjøre og.. Hvor de trenger hjelp, hvor de står fast.* Den kvinnelige læreren på skole G (G1) som hadde tatt flest kurs i vurdering, var også opptatt av å lære mer vurderingsmetodikk: - *Jeg har hatt kursing i mappevurdering og andre vurderingsformer, så jeg føler at jeg kjenner til mange metoder. Men det jeg kunne tenkt meg og som jeg føler at jeg trenger, er kanskje mer inngående opplæring eller enda mer praktisk opplæring i mappevurdering. For jeg kan alle de overordnede teoriene, men det å gjennomføre det i praksis er mye vanskeligere så det kunne jeg godt tenkt meg mer veiledning om.*

Svar fra disse intervjuene og intervjuene fra KUPP prosjektet, Klettets rapport og analysen av Elevinspektørene 2005, kan tyde på at den norske skolen og lærerutdanningen har en utfordring i å utforme en pedagogikk der vurdering inngår i mer systematiske opplegg og læringsprosesser. Metodisk mangfold er et av nøkkelbegrepene som knyttes til profesjonsbegrepet i KUPP prosjektet (jf Hoel 2005) og mine intervjuer og intervjuene i KUPP prosjektet kan tyde på at den norske læreutdanningen kommer til kort når det gjelder dette punktet. Fagdidaktikken i lærerutdanningen tar for seg vurdering ved at fagdidaktikk omfatter spesifisering av mål, innhold, metode og vurdering i forhold til de enkelte skolefaglige aktiviteter (Dale 1999). Likevel tar den ikke opp relasjonen mellom skolefagene og heller ikke opp vurderingsmetoder og kriterier som kan overføres i ulike fag. Dette gjøres i Alberta der man har tilrettelagt for et grunnlag for å bruke vurdering som profesjonsfaglig kompetanse med fokus på kunnskap om vurdering og vurderingsmetoder.

7.2 Hva kan vi lære av erfaringene fra skole A og B?

I følge Le Métais (1999) er ethvert utdanningssystem en kombinasjon av fortiden, nåtiden og fremtiden. Fortiden er representert gjennom mål, verdier og strukturer og læremetoder. Nåtiden er representert i pågående vurderingsopplegg og prosjekter som har en umiddelbar påvirkning på læreplanpraksisen i skolen. Fremtiden er representert i forhold til læreplanen og som læringsmål (Le Métais 1999). Med erfaring fra ulike fortid, utdanningspolitisk utgangspunkt og læreplantradisjon har vurderingskulturen i Alberta og Norge utviklet seg ulikt. Alberta har vært preget av nyliberale strategier som vektlegger ”accountability” og effektivitet i over 20 år, og dette har preget deres behov for vurdering. Videre har Canada vært preget av den amerikanske læreplantradisjonen som er metodestyrte i stor grad (jf Gundem og Hopman 2002). Norge derimot, der didaktikken har stått sterkt, har ikke blitt påvirket av nyliberalismen i like stor grad. Videre har enhetsskolen vært et varemerke for norsk skole helt fram til slutten av 90 – tallet (jf Telhaug 1994).

Til tross for ulikt utdanningspolitisk utgangspunkt tar Norge nå i bruk lignende vurderingsopplegg og strategier (nasjonale / provinsielle prøver) som Alberta har hatt i siden tidlig på 80 - tallet for å nå den nye læreplanens mål – et kunnskapsløft. Alberta som har nådd dette målet (jf PISA resultater) har lang erfaring med vurdering og jeg synes det viktig å ta tak i disse erfaringene. Alberta har opplevd både positive og negative konsekvenser med ”vurdering AV læring” i sin provins. På bakgrunn av mine empiriske undersøkelser i de to barneskolene i Edmonton har jeg kategorisert erfaringene slik:

- *Vurdering AV og FOR læring behøver ikke sees på som å to motsetninger. I stedet kan de sammen bidra til læring (jf skole B og Black og Wiliam 1998b).*
- *Tydeliggjøring av læringsmål og vurderingskriterier kan bidra til å inkludere elevene i læringsprosessen (jf Black og Wiliam). Videre kan kriterier være en hjelp for læreren og eleven selv til å finne elevens ståsted (jf lærere på skole B).*
- *Vurdering AV læring kan føre til utenatføring og konkurranse dersom man overfokuserer på resultatene (jf tendenser i skole A).*
- *Lærerutdanningen bør fokusere på vurdering som en selvstendig fagdisiplin for å skape et grunnlag for å bruke vurdering som profesjonsfaglig kompetanse og som verktøy for læring.*

- *"Skolekoden" styrer lærernes vurderingspraksis i større grad enn læreplanen og nasjonale prøver ved at man tolker bruken av læreplanen og prøvene ut fra skolens behov som preges av skolens filosofi, ledelse og fastsatte rammer.*
- *En innføring av nasjonale prøver behøver ikke nødvendigvis svekke lærerprofesjonaliteten. Støtteapparatet rundt og rammebetingelsene er avgjørende for hvorvidt læreren har mulighet til å bruke vurdering som profesjonsfaglig kompetanse eller ikke (Jf Edmonton Public School Board).*

Disse punktene er en oppsummering av erfaringer gjort i skole A og B i Edmonton. Jeg vil nå utdype disse punktene og relatere de til den norske vurderingsdebatten.

7.2.1 Samsvar mellom summativ og formativ vurdering

- *Vurdering AV og FOR læring behøver ikke sees på som å to motsetninger. I stedet kan de sammen bidra til læring (jf skole B, Black og Wiliam 1998b).*

To av formålene Utdanningsdirektoratet (2006) beskrev i forbindelse med innføring av nasjonale vurderingskriterier var å "sikre sammenhengen mellom underveis - og sluttvurdering" og "bidra til en likeverdig og rettferdig vurdering av elevene" (jf del 1.2). Som jeg var inne på tidligere, har mye av tilbakemeldingen i barneskolen i Norge vært opp til den enkelte lærer fordi man ikke har hatt noen formelle vurderingsstrategier. Den uformelle formative vurderingen har stått i motsetning til den formelle vurderingen ved at man har laget et skille mellom individrelatert vurdering på den ene siden og målrelatert vurdering på den andre. Nasjonale prøver og den daglige elevvurderingen representerer både summativ og formativ vurdering med to ulike formål. Den formative skal bidra til læring og utvikling, mens den summative skal gi en oversikt over resultat og produkt. Som jeg beskrev i analysen, fokuserte lærerne i skole B like mye på formative vurderingsstrategier som summative vurderingsstrategier og de knyttet disse sammen ved hjelp av tydelige kriterier. Lærerne mente at det var urealistisk at elevene kan praktisere et slikt skille og de så på vurdering som mer enn bare en prøve. I følge Black og Wiliam (1998b) bør man kombinere summative og formative vurderingsformer for å minske skillet både for lærere og elevene. Et eksempel på dette er å gi elevene en oppgave som andre medelever kan gi respons på eller som elevene selv vurderer ved hjelp av egenvurderingsskjemaer. Kriterier spiller da en viktig rolle for at elevene skal vite hva de skal kunne for å nå et læringsmål.

7.2.2 Tydelighet og kriteriesetting

- Tydeliggjøring av læringsmål og vurderingskriterier kan bidra til å inkludere elevene i læringsprosessen (jf Black og Wiliam). Videre kan kriterier være en hjelp for læreren og eleven selv til å finne elevens ståsted (jf lærere på skole B).

Et annet formål for nasjonale vurderingskriterier som Utdanningsdirektoratet (2006) har beskrevet er å ”styrke vurdering av den enkelte elev underveis gjennom presise faglige tilbakemeldinger” (jf del 1.2). Nasjonale undersøkelser som jeg har henvist til tidligere i oppgaven tyder på at vurderingen især på småskoletrinnet er preget av lite systematikk og tydelighet. Formative vurderingsstrategier som gir eleven tydelige tilbakemeldinger underveis i læringsprosessen kan bidra til selvregulert læring, økte forventninger til egen mestring (jf Bandura i Pervin & John 2001: 449) og igjen økt læringsutbytte. Dette er i tråd med Black og Wiliam (1998b) som mener at synliggjøring av kriterier og læringsmål kan bidra til effektiv læring som legger vekt på elevens forståelse i læringsprosessen.

Alle de fire offentlige skolene som var styrt av EPSB jobbet med spesifisering av læringsmål og kriteriesetting for å inkludere eleven mer i læringsprosessen. Påstanden over er i tråd med hva de tre lærerne på skole B mente var det mest positive med å tydeliggjøre læringsmål og vurderingskriterier. Læringsmålene ble som regel hentet fra læreplanen, mens vurderingskriteriene kunne lærerne lage selv i eller sammen med elevene. I den norske vurderingsdebatten har man i større grad fokusert på de negative konsekvensene vurderingskriterier kan føre til som blant annet at det kan sees på som ”bedømming” av elevene (jf eksempel fra Dagsavisen s. 19). Videre har man vært engstelig for at kriterier vil være ødeleggende for de svakeste elevene (jf ”Monahistorien” s. 29). I følge Black og Wiliam (1998b) er ”vurdering FOR læring” i form av egenvurdering og tydelig kriteriesetting spesielt viktig for de svake elevene og har en tydelig effekt i deres læringsprosess. I skole E der ca 45 % av elevene hadde spesielle vansker var egenvurderingsskjemaer mye brukt. Elevene vurderte seg selv opp mot kriterier som de selv hadde satt i felleskap med de andre elevene eller læreren. Økt fokus på kriteriesetting i barneskolen vil også kunne føre til at overgangen til ungdomsskolen ikke vil bli så vanskelig for de svakeste elevene.

I alle de fire skolene som var styrt av EPSB jobbet de mye med formative vurderingsmetoder og kriteriesetting og dette var satt i system. Metodene lærerne brukte for å gjøre dette på var ulike, men tankegangen var den samme. Ved å gi tydelige tilbakemeldinger på grunnlag av kriterier fikk de elevene med i vurderingsprosessen. Videre ble elevene i større grad stilt til ansvar for sin egen læringsprosess. Jeg vil her henwise til et eksempel fra Osloskolen, der det allerede er gjort et forsøk på å overføre en slik tankegang ved hjelp av kriterier. I prosjektet ”Dialogbasert kvalitetsutvikling i skolen – en ledelsesutfordring” samarbeidet Oslo kommune med Utbildningsförvaltningen i Stockholm. Sammen gjorde de et pilotprosjekt der de innførte LUS (Lese Utviklings Skjema) i flere Osloskoler. LUS er et verktøy for vurdering av elevers leseutvikling, mål- og resultatstyring der man følger elevens leseutvikling ved hjelp av kriterier. Dette verktøyet har hjulpet flere lærere i barneskolen med å se hvilket nivå elevene er på. Mange elever som ikke kan lese, får heller ikke med seg innholdet i andre fag og dette skjemaet blir derfor nyttig også for andre fag. LUS skjemaet gjør det mulig å tydeliggjøre leseprosessen ved hjelp av kriterier som må nås for å nå visse stadier. Samtidig som det er et verktøy for lærerne, er det også et viktig redskap for elevene selv i sin egenvurderingsprosess. Skjemaet har gjort det lettere for både lærere og elever å sette ord på hva de kan og hva de må jobbe bedre med (Allgard, Rudquist & Sundblad 2006). Med andre ord er ikke en slik tenkning der kriterier brukes for å inkludere både lærer og elev i læringsprosessen fremmed for mange norske lærere.

7.2.3 Vurdering kan føre til konkurranse og utenatføring

- *”Vurdering AV læring” kan føre til konkurranse og utenatføring dersom man overfokuserer på resultatene.*

Et dilemma Utdanningsdirektoratet (2006) beskrev vedrørende en innføring av nasjonale vurderingskriterier, var at det kunne føre til ”en konkurrerende læreplan dersom kriteriene vektlegger bare noen av kompetansemålene” (jf del 1.2). Erfaringene fra skole A viste at en overfokusering av ”vurdering AV læring” kunne føre til konkurranse. Fokuset på ”vurdering AV læring” var stort og det var dette de brukte som mål på elevenes læring. Videre hadde de laget sin egen læreplan der de vektla bare noen av kompetansemålene fra læreplanen i Alberta. Lærere på skolen fortalte meg at de gjorde dette for at målene lettere kunne vurderes. Alle elevene fikk prøver i alle fag hver uke med karakterer. Dette gjorde at konkurransen om karakterene mellom elevene og klassene seg i mellom var stor.

Skole A er et eksempel på en skole som kan fremme utenatføring og konkurranse fordi resultatene kommer foran elevenes behov. Videre var denne skolen rangert høyt oppe på listene blant alle de andre skolene i Alberta, noe lærere og elever var stolte av. Dette kan tyde på at de provinsielle testene ikke måler elevenes egentlige forståelse og dybdekunnskap, men at man kan få gode resultater ved hjelp av utenatføring. Det er nok ikke usannsynlig at flere skoler i Alberta kan ha vært preget av en slik vurderingskultur. Den Canadiske analysen av PISA rapporten 2003 viser at elevene i Alberta ikke bare skårer høyest i Canada når det gjelder faglige ferdigheter. De skårer også høyest i Canada når det gjaldt utenatstrategier. Videre har Alberta "highest level of preference of competitive learning situations" (Bussiere, Cartwright og Knighton 2004:49).

Det hadde vært interessant å se hvor skole A og B plasserte seg i forhold til hverandre på disse områdene fordi erfaringene tyder på at skole A vil kunne ha høyere grad av utenatføring og mer konkurranse enn skole B. Jeg har dessverre ingen tall på dette fordi ingen av de jeg snakket med hadde en så detaljert oversikt over skolenes plassering sett i forhold til Albertas PISA resultater.

7.2.4 Vurdering som en selvstendig fagdisiplin

- Lærerutdanningen bør fokusere på vurdering som en selvstendig fagdisiplin for å skape et grunnlag for å bruke vurdering som profesjonsfaglig kompetanse og som verktøy for læring.

I henhold til Black og Wiliam (1998) skjer læring først og fremst innenfor de fire veggene i klasserommet og det er behov for å bedre de metoder som brukes der. Mine intervjuer i de to Osloskolene og svarene i KUPP prosjektet tyder på at det er et behov for kunnskap om flere vurderingsmetoder og kunnskap om hvordan vurdering utføres i praksis. Nye vurderingsformer og arbeidsmåter krever ny kunnskap og her har lærerutdanningen en viktig rolle når det gjelder å gi lærerstudenter grunnleggende kunnskap som de kan tilpasse sin undervisningspraksis. I lærerutdanningen i Norge er vurderingen mer en del av fagdidaktikken og ikke en selvstendig fagdisiplin.

Lærerutdanningen i Alberta hadde et obligatorisk vurderingskurs i lærerutdanningen der man knyttet vurdering til praksis, og til alle de ulike fagene som lærerne skulle undervise i og ga lærerne et stort utvalg av både formative og summative vurderingsmetoder. Som jeg har vært inne på tidligere i oppgaven knyttes praksisnærhet, tverrfaglighet og metodisk mangfold alle

nøkkelbegrep til profesjonsbegrepet og profesjonsfaglig kompetanse (jf Hoel 2005). Ut fra dette kan man si at lærerutdanningen i Alberta har et profesjonsaspekt og gir et grunnlag for å bruke vurdering som profesjonsfaglig kompetanse. Til tross for samme lærerutdannelse, har lærerne på skole A og B utviklet seg forskjellig i forskjellig retning. I skole B var det rom for undersøkende pedagogikk som aksjonsforskning og utprøving av nye metoder fordi det var et ønske fra EPSB at lærerne jobber med å utvikle seg profesjonelt. Skole A er et eksempel på en skole der man ikke undersøkte praksis og metoder, men fulgte de metoder som var prioritert på skolen. Denne skolen verdsetter "accountability" i stor grad ved at både elever og lærere og skolen som helhet ansvarliggjør seg ved å vise til sluttresultater.

7.2.5 Ledelsen bør delta i vurderingsarbeidet

- "Skolekoden" styrer lærernes vurderingspraksis i større grad enn læreplanen og nasjonale prøver ved at man tolker bruken læreplanen og prøvene ut fra skolens behov som preges av skolens filosofi, ledelse og fastsatte rammer.

At ledelsen må delta for å skape en kultur for vurdering som skaper læring er i henhold til Øygarden (2006) som deltok i prosjektet "Vurdering som profesjonsfaglig kompetanse". Han mener det er viktig å løfte vurderingsaktiviteten fram fra å representere en individuell forståelse og praksis hos den enkelte lærer til å bli et felles anliggende for skolen og kollegiet. I et slikt perspektiv er det avgjørende at ledelsen ser på vurdering som ett av skolens prioriterte områder. Øygarden mener at det må legges organisatorisk til rette for at vurderingen finner sted på en hensiktsmessig måte, det må være en del av lærersamarbeidet å utvikle felles forståelse av krav og forventninger til elevenes arbeider på ulike nivå, blant annet må det være et tolkningsfellesskap som skaper trygghet for at vurderingen er noenlunde valid. Å utvikle en slik felles tolkingskompetanse vil sikre at skåring av nasjonale prøver, bruken av ulike ikke-standardiserte prøver og den generelle faglige vurderingen på skolen blir mer reliabel (Øygarden 2006). Dette er et viktig poeng og erfaringene fra skolene i EPSB viser at lærerne har mulighet for dette blant annet ved å delta i "Professional learning communities".

De siste årene har man også i Norge vært opptatt av at man jobber sammen på flere nivå for en felles kultur som verdsetter læring i skolen og vurdering har blitt knyttet opp mot dette. I "En snarvei til Kompetanseberetningen for Norge" (2005) med navnet "Lærer elevene mer på lærende skoler?" (UFD) er målet å gå skape lærende skoler i Norge ved å stimulere til en

målrettet utvikling. Med denne beretningen har man gjort et første forsøk på om det er sammenheng mellom skolens resultater og satsning på organisasjonsutvikling. Det blir henvist til en undersøkelse fra Læringslaben som tyder på at det er en sammenheng mellom organisasjonsutvikling, samhandling og forventninger i det voksne miljøet i skolen, og elevenes læring (UFD 2005: 65). For å skape lærende skoler blir tydelighet vektlagt som en viktig forutsetning. I Stortingsmelding nr 30 står det skrevet om lærende organisasjoner at ”I lærende organisasjoner er forventningene og tilbakemeldingene tydelige. Lærende organisasjoner stiller derfor særlig store krav til tydelig og kraftfullt lederskap som er seg bevisst skolens kunnskapsmål” (St. meld 30: 27). Med andre ord er det et ønske fra politisk hold å skape lærende skoler i Norge ved hjelp av tydelige forventninger og tilbakemeldinger og fokus på målrettethet. I EPSB var de opptatt av å skape lærende skolemiljøer, og ”professional learning communities” var et middel til å oppnå dette. Som jeg var inne på i del 6. 1.3 er det vanskelig å oppnå lærende organisasjoner i praksis blant annet fordi andre motiver enn læring vil være styrende. Likevel tyder erfaringene fra EPSB på at det er mulig å lære på flere nivå dersom man har en felles visjon og dersom forholdene legges til rette for at visjonen kan følges. I EPSB jobbet man i samarbeid med blant annet ”Alberta Assessment Consortium” for å oppnå ”assessment FOR learning”. Intervjuene fra skole B og observasjonene fra skole B, C, D og E tyder på at denne visjonen ble utført i praksis ved hjelp av ”Professional learning communities”.

7.2.6 Vurdering behøver ikke nødvendigvis bety mistillit til lærerprofesjonaliteten

- En innføring av nasjonale prøver behøver ikke nødvendigvis svekke lærerprofesjonaliteten. Det politiske støtteapparatet rundt, ledelsen og rammebetingelser er avgjørende for hvorvidt læreren kan bruke vurdering som profesjonsfaglig kompetanse (jf Edmonton Public School Board og skole B).

Utdanningsdirektoratet (2006) påpeker som en negativ konsekvens at nasjonale vurderingskriterier kan føre til ”en mistillitserklæring til lærerprofesjonaliteten” (jf del 1.2). Ut fra min empiriske undersøkelse syntes ikke dette å være tilfelle i noen av de to skolene jeg undersøkte. Erfaringene fra skole B viste at det er mulig å jobbe som profesjonell lærer, dersom systemet og rammene man jobber i tillater det. Videre er det viktig at hele kollegiet er med på en slik tenkning. Erfaringene fra skole B, viser at det er mulig å sette profesjonelle strategier i system ved å gi lærere rammer som gir dem større handlingsrom og mulighet til metodemangfold. Lærerne brukte varierte formative og summative vurderingsformer, noe

som gjorde at de brukte vurdering som profesjonsfaglig kompetanse i større grad i skole B enn skole A.

De siste årene har mye endret seg i den norske utdanningspolitikken og ”accountability - bølgen” har nådd Norge. Som jeg har vært inne på tidligere (jf del 2.5) skiller Broadfoot (1996) mellom vurderingens tre roller som er 1) vurdering som hjelp for lærer og elev, 2) vurderingens rolle i læreplanen og 3) vurdering som et politisk verktøy for kontroll. I dagens globale samfunn der kunnskap og kompetanse sees på som et økonomisk gode, kan man ikke lenger komme vekk fra vurderingens kontrollfunksjon. Dette er i henhold til Lundgren som på 1990 – tallet påviste at vurderingssystemet har en sterk styrende innvirkning på det som skjer i opplæringen (Lundgren i Engelsen 2006: 245). I EPSB har man på et vis utnyttet vurderingens kontrollerende side til å motvirke konsekvensene resultatorienteringen har ført til. Ved å gi rammer som åpner for ”vurdering FOR læring” og ”professional learning communities” blir læring satt i fokus og vurdering blir brukt med pedagogikken som grunnlag (jf Dale 1999). Erfaringene fra EPSB tyder på at Broadfoots tre punkter vedrørende vurderingens roller ble knyttet tett sammen. For det første var det et stort fokus på formative vurderingsmetoder der vurdering var en hjelp for både lærer og elev. For det andre var disse metodene nedfelt i læreplanen. For det tredje var det lagt politiske føringer som støttet formative vurderingsstrategier både fra EPSB og Alberta Education.

På bakgrunn erfaringene i Edmonton har jeg laget en modell (vedlegg 3) som viser hvordan vi i Norge kan legge til rette for at lærere kan bruke vurdering som profesjonsfaglig kompetanse og som verktøy for læring innenfor de politiske rammer som har satt kriterier for hva som er god kvalitet i norsk skole (jf nasjonalt vurderingssystem).

8 Avsluttende kommentar

Så hva er Albertas hemmelighet? Hvorfor skårer de så høyt i PISA? Som jeg har vært inne på tidligere er læreplanen i Alberta detaljert i stor grad, men det er også mange muligheter for kreativitet og tolkning. Jeg vil ikke gi noe entydig svar på om Alberta skårer høyt i PISA undersøkelsen på grunn av deres lærerutdanning, vurderingsfokus eller deres læreplan. Likevel tyder erfaringene fra Edmonton på at vurderingskulturen i Alberta har preget lærerprofesjonen i stor grad, både på godt og vondt. Grunnet det store fokuset på ”ansvarliggjøring” og vurdering stilles det høye krav til både elever, rektorer og lærere noe som sannsynligvis vil få konsekvenser for praksisen i skolen.

”Alberta Education” (2005) tror selv at deres suksess i PISA og nasjonalt er deres store fokus på vurdering og at deres provinsielle prøver er tett knyttet opp mot læreplanen. Dette kan være et viktig poeng fordi lærere og elever er vant til prøver og det er en tydelig sammenheng mellom det som skal læres og det som skal vurderes. Likevel betyr ikke en slik tett sammenheng nødvendigvis at det fører til læringsprosesser som fremmer forståelse. I følge Black og Wiliam (1998b) kan tydelige vurderingskriterier og læreplanmål føre til to ulike innlæringsmåter som jeg har vært inne på tidligere. Den første som de kaller *ineffektiv læring*, preges av en ”teaching to the test” undervisning der man legger opp til å lære flest mulig læringsmål. Den andre innlæringsmåten som de kaller *effektiv læring* fremmer forståelse i læreprosesser ved at eleven selv blir inkludert i læringsmålene på ulike måter. Albertas høye skåre når det gjelder utenatstrategier i PISA 2003 (jf del 7.2.3) kan tyde på at flere skoler enn skole A preges av ineffektiv læring der bare noen av kompetansemålene vektlegges. PISA undersøkelsen er resultatbasert og vil derfor ikke vise elevenes dybdeforståelse. Med andre ord kan det være mulig å oppnå gode resultater ved hjelp av utenatlæring. Skole A er et eksempel på dette. Jeg vil igjen henvise til lærer A3 som sa: - *you can be worried and that is ok with the superficial kinds of things that you need to learn.* Med andre ord kan man lære det overfladiske innholdet som *må* læres til tross for resultatpress og angst for prøver. Skole A var en av de skolene som lå blant de 30 beste i Alberta provinsen i 2005, noe som kan tyde på at ikke heller de provinsielle prøvene nødvendigvis kartlegger elevenes helhetlige forståelse. I tilknytning til dette er det viktig å tenke på at det ikke er de nasjonale eller provinsielle prøvene i seg selv som skaper elever med høyt kunnskapsnivå. Det er bruken av dem. Erfaringene fra skole A og B viser hvordan

”vurdering AV læring” kan brukes på to ulike måter. 1) som kartlegging med formål om å forbedre læring (jf skole B) og 2) kartlegging med formål å vise til resultater (jf skole A).

Jeg vil avslutte med å gi en oppsummering av hvordan preger Albertas vurderingskultur læreres bruk av vurdering som profesjonsfaglig kompetanse og som verktøy for læring i skole A og B (jf problemstilling nr. 1). Som jeg allerede har vært inne på tyder erfaringene fra skole B på at en innføring av nasjonale prøver ikke nødvendigvis behøver å svekke lærerprofesjonaliteten. Sammenligningen av skole A og B viser at ”skolekoden” styrer lærernes handlingsrom i større grad enn læreplanen og nasjonale prøver ved at ledelsen og lærerne tolker læreplanen og bruken av de provinsielle prøvene ut fra skolens filosofi, ledelse og fastsatte rammer. For å støtte læreres profesjonelle kompetanse og varierte vurderingsmetoder la EPSB rammer for rektorer og lærere som gjorde det mulig å opprettholde gode pedagogiske metoder for elevmedvirkning ved hjelp av fokus på kriteriesetting og tydeliggjøring av læreplanmål.

En undersøkelse som støtter dette ble utført i januar 2006 av forsker Åsmund Arup Seip ved Fafo³¹ på oppdrag Utdanningsforbundet. I denne undersøkelsen blir arbeidsvilkårene for hvordan lærere blir regulert i sentrale avtaler i Norge, Danmark, Sverige og Finland sammenlignet. I en kommentar understreker leder for Utdanningsforbundet Helga Hjetland at – *De finske lærerne har lang og god utdanning, og de har stort handlingsrom i forhold til å legge til rette for god undervisning*. Hun påpeker at det er en tydelig sammenheng mellom elevresultater, lærerkompetanse og lærernes muligheter til å legge til rette for god undervisning³². Dette er i tråd med erfaringene fra EPSB som tydet på at det ble lagt rammer som ga lærerne mulighet til å tilrettelegge god undervisning og gode vurderingsformer. Som erfaringene fra EPSB har vist er det mulig å legge politiske rammer som kan tilrettelegge for læringsfremmede metoder. Ved hjelp av kunnskap om hvordan bruke vurdering som profesjonsfaglig kompetanse og som verktøy for læring fikk lærere i EPSB mulighet til å bruke vurderingsmetoder som fremmer og ikke hemmer læring.

Erfaringene fra Edmonton kan være med på å berike vurderingsdebatten i Norge på flere måter (jf problemstilling nr. 2). Som jeg har redegjort for i del 7.2 kan ”vurdering AV læring

³¹ Norsk forskningsstiftelse

³² http://www.utdanningsforbundet.no/UdfTemplates/Page____30737.aspx

og fokus på vurderingskriterier føre til både negative og positive konsekvenser. Ved å ta utgangspunkt i erfaringer fra skole B har jeg gitt et eksempel på at vurdering FOR og AV brukes forskjellige og ikke behøver å sees på som å to motsetninger. I stedet kan de sammen bidra til læring (jf Black og Wiliam 1998b). I forbindelse med Finlands suksess i PISA undersøkelsen sier en finsk professor med navnet Hämäläinen: ”Evaluerer er som alkohol – å drikke litt er sunt, men for mye er bare skadelig” (”Bedre Skole” 2005, nr 4: 7). Hämäläinen bruker her begrepet ”evaluerer”, om det jeg i denne oppgaven har kalt ”vurdering AV læring”. Likevel passer dette sitatet godt inn i forhold til skillet mellom skole A og B som begge preges av provinsielle prøver og vurderingskulturen i Alberta. Skole A har latt disse prøvene styre vurderingspraksisen i skolen ved å overfokusere på ”vurdering AV læring” som vurderingsform. Skole B derimot bruker ”vurdering AV læring” med måte ved å vektlegge ”vurdering FOR læring” i tillegg. Om de kommende nasjonale prøvene i Norge vil skape skoler som ligner skole A eller skole B er derfor avhengig av hvordan vi vil bruke vurderingen som verktøy for læring og hvordan det utdanningspolitisk vil bli tilrettelagt for dette.

Vedlegg 1: funn fra PISA 2000 / 2003

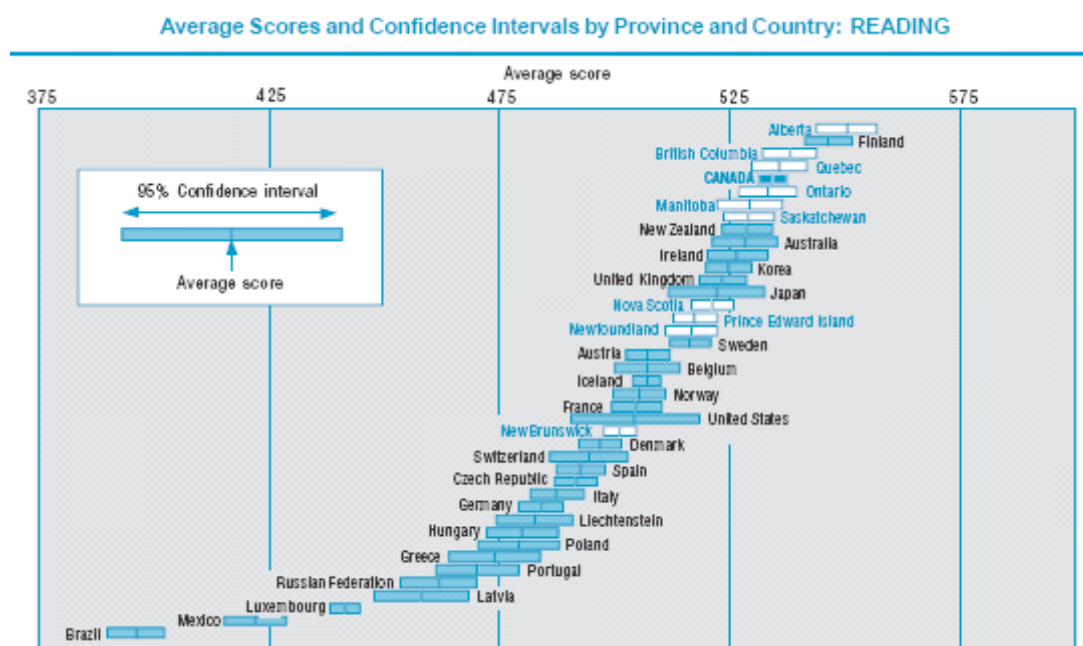
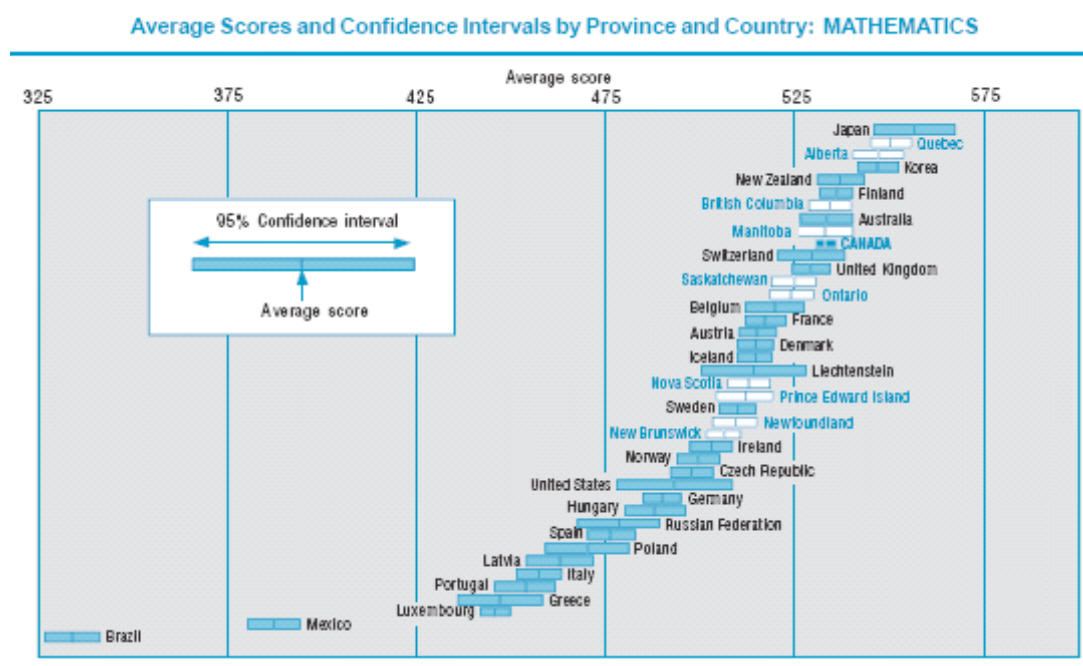


FIGURE 1.2



Figur 3 og 4 : PISA 2000 (Bussiere med fler 2001)

Average Scores and Confidence Intervals for Provinces and Countries: READING REFLECTING

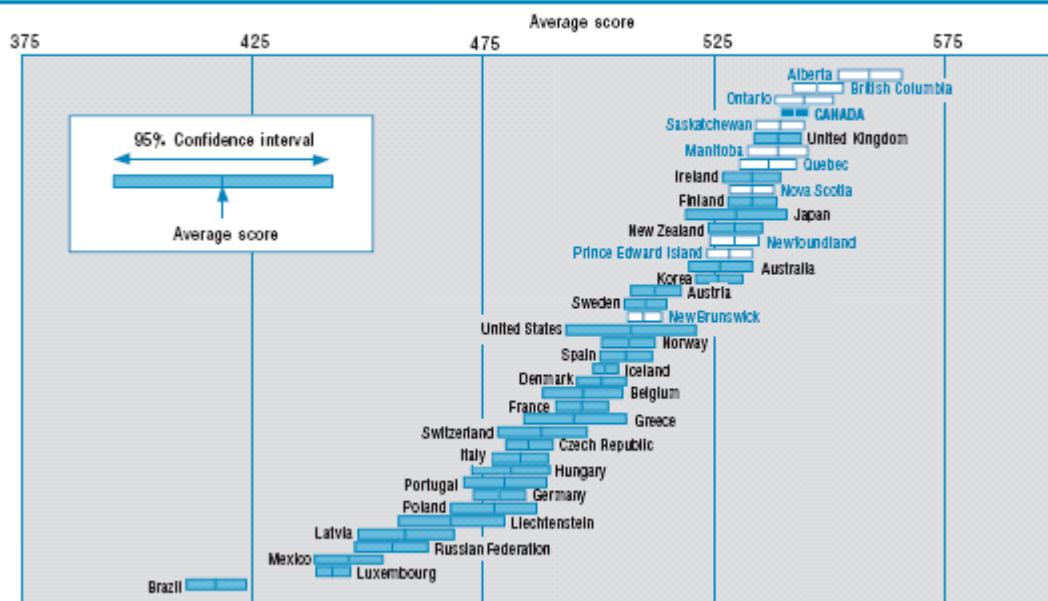
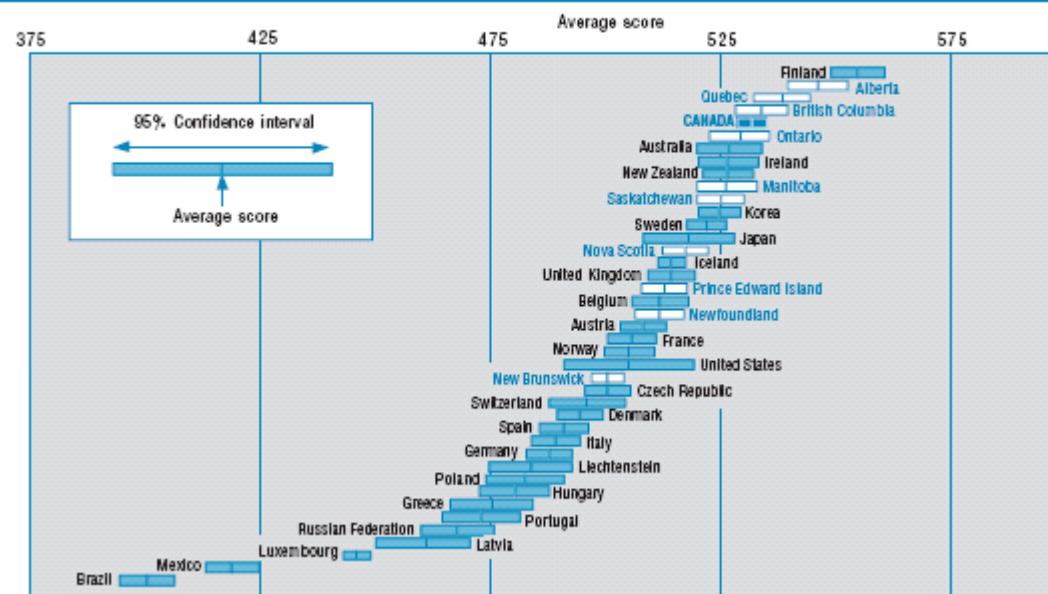


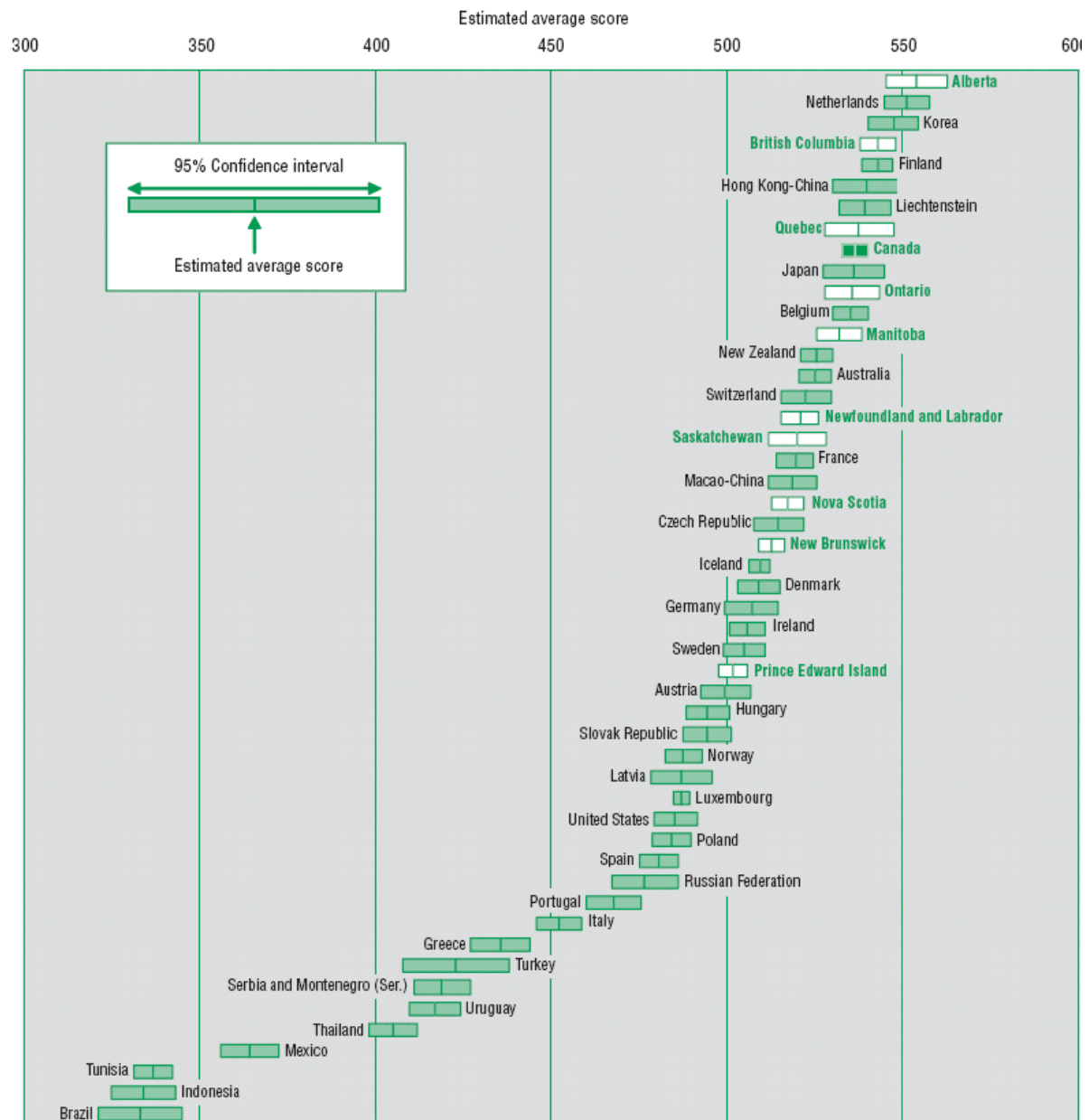
FIGURE 1.6

Average Scores and Confidence Intervals for Provinces and Countries: READING INTERPRETING



Figur 5 og 6 PISA: 2000 (Bussiere med fler 2001:14)

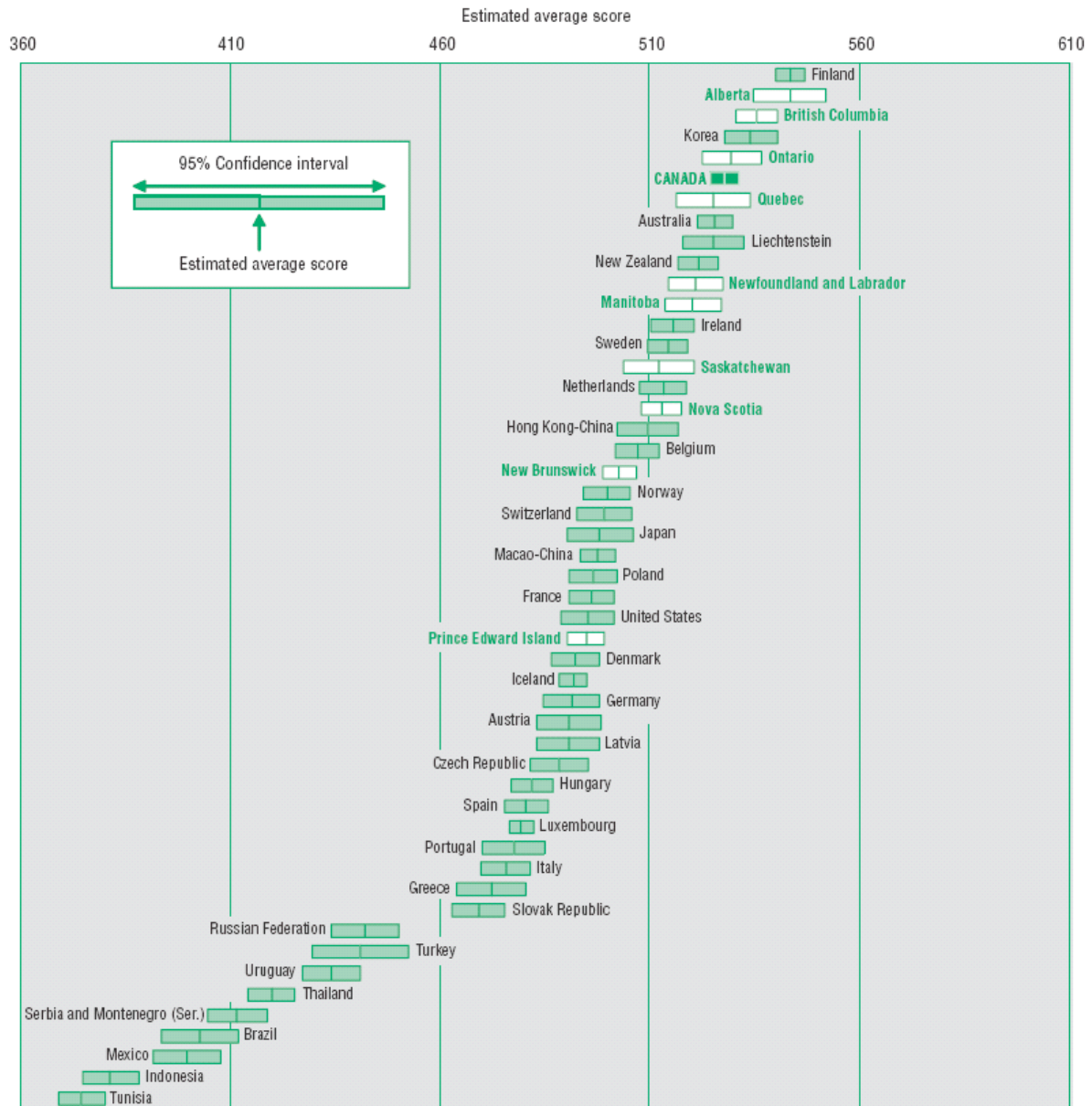
Estimated average scores and confidence intervals for provinces and countries:
MATHEMATICS *change and relationships*



Note: The OECD average is 499 with a standard error of 0.7.

Figur 7: PISA 2003 (Bussiere, Cartwright & Knighton 2004:18)

Estimated average scores and confidence intervals for provinces and countries: READING



Note: The reading results for 2003 are based on the reading literacy proficiency scale that was developed for PISA 2000 which had a mean of 500 for the 27 countries that participated in PISA 2000. However, because three additional OECD countries are included in the PISA 2003 reading test, the overall OECD mean for PISA 2003 is 494 with a standard error of 0.6.

Figur 8: PISA 2003 (Bussiere, Cartwright & Knighton 2004:26)

Vedlegg 2: intervjuguide

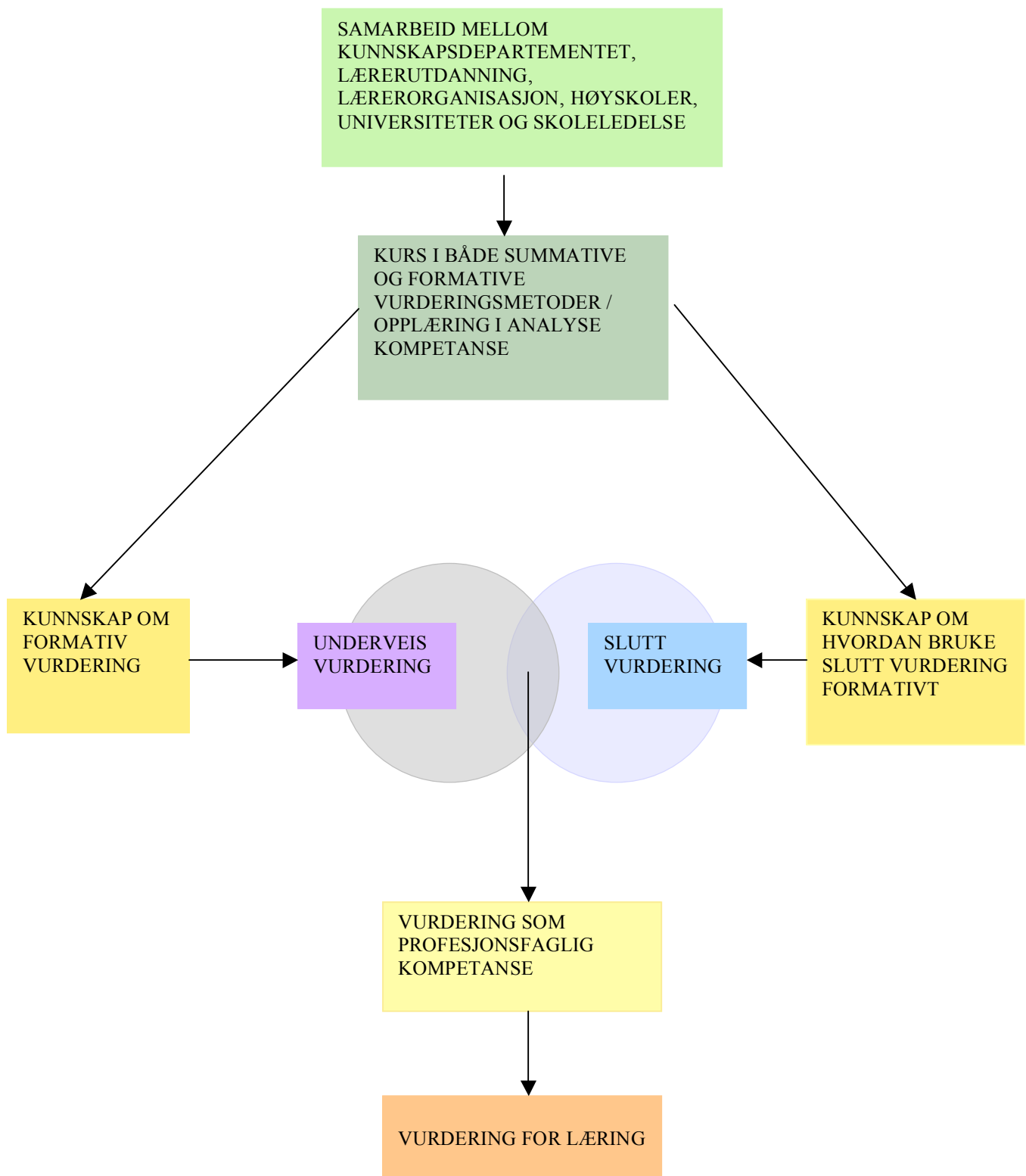
INTERVJU GUIDE (engelsk versjon)

1. Age
2. For how long time have you worked as a teacher
3. What kind of education do you have?
4. Would like to be even more educated? Why / why not?
5. Why did you become a teacher?
6. What is being assessed when you do student assessment?
 - Working effort
 - The student's ability to plan his or her own schoolwork
 - Knowledge, the result, the product
 - The process
 - Social skills
 - Attitude
7. What is the aim of the student assessment?
8. What do you use the student assessment for? / What kind of function does it have?
9. What kind of role does the student play in the assessment process?
10. Who is making the criteria for the student assessment?
11. What do you think is the advantage of doing a lot of student assessment?
12. Have you taken any courses on assessment?
13. Where did you learn the methods and strategies that you use?
14. Do you feel that you need more knowledge about assessment strategies?
15. What do you know about your colleague's assessment strategies?
 - Are there big differences in your school on how the assess?
 - Do you have some common rules or strategies that you use?
16. How do you inform the parents about how their children achieve?
17. Do feel that you student's parents have great knowledge about their children's abilities and results?
18. Do you feel that the parents have some influence on what is happening in the school?

INTERVJU GUIDE (norsk versjon)

1. Alder
2. Hvor lenge har du jobbet som lærer?
3. Hva slags utdanning har du?
4. Kunne du tenke deg å ta noen form for videreutdanning?
5. Hvorfor ble du lærer?
6. Hva blir vurdert når du vurderer en og en elev.
 - Arbeidsinnsats, evne til å planlegge eget arbeid
 - Kunnskaper, faglig resultat, produkt
 - Det sosiale, holdninger
 - Prosessen
7. Hva slags funksjon har vurderingen?
8. Hva bruker du vurderingen til?
9. Hva slags rolle har eleven i vurderingsprosessen?
10. Hvem avgjør utviklingen av kriterier for elevvurdering?
11. Føler du behov for mer kunnskap om vurderings strategier og retningslinjer for hvordan vurdere elevene dine? Hvorfor / hvorfor ikke?
12. Har du tatt noen kurs som handler om elevvurdering?
13. Hvor har du lært de strategier du bruker?
14. Hva lærte du om vurdering i lærerutdanningen?
15. Hvordan føler du at du selv ble vurdert da du gikk på lærerutdanningen?
16. Hva tenker du er målet med den vurderingen som du gir dine elever?
17. Hva vet du om dine kollegers vurderingspraksis?
 - Er det stor forskjell på hvordan dere vurderer elevene innad i skolen?
 - Har dere noen felles retningslinjer for hvordan dere skal vurdere?
18. Det skal fra høsten 2007 innføres skriftlig vurdering på barnetrinnet i Oslo skolene?
Hva tenker du om det? Hvilke positive og negative sider tror du at dette kan føre til?
19. Har du nå et godt samarbeid med foreldrene til dine elever?
20. I hvilken grad føler du at foreldrene har kunnskap om elevenes ferdigheter på skolen?

Vedlegg 3: Modell



Referanser

Alberta Assessment Consortium (2006): *Refocus: Looking at assessment for learning. Second Edition*. Alberta: AAC

Alberta Education (2005): *National & International Studies of Achievement*

Alberta Teachers Association (2006) *Handle With Care: Futures being built. An Invitational Symposium on Accountability in Education*. Alberta: ATA

Alberta Teachers Association (2002): *A Brief History of Public Education in Alberta*. Alberta: ATA

Allard, Birgita, Rudquist & Sundblad, Bo (2006): *Den nye LUS boken. En bok om Leseutvikling*. Oslo: Cappelen Akademiske Forlag

Bedre skole (2005): nr 4. *Tidsskrift for lærere og skoleledere*.

Black, Paul & Wiliam, Dylan (1998 a): *Inside the Black Box – Raising standards through classroom assessment*. London: School of Education, King's college.

Black, Paul & Wiliam, Dylan (1998 b): *Working inside the Black Box - Assessment for learning in the classroom*. London: School of Education, King's college

Broadfoot, Osborn, McNess Panel, Ravn, Triggs (2003): *A world of Difference. Comparing Learning across Europe*. England: Open University Press

Broadfoot, Patricia. (1996): *Assessment and learning: Power or Partnership?* Chichester: John Wiley & Sons. In Goldstein, Harvey and Lewis Toby (red.): *Assessment; Problems, Developments and Statistical Issues*, 21–40.

Bussiere, P. Cartwright, F. Crocker, R. Ma, X. Oderkrik, J. & Zhang, Y (2001): *Measuring up. The Performance of Canada's Youth in Reading, Mathematics and Science OECD PISA study – First results of Canadians aged 15*. Ottawa, ON: Statistics Canada. Se: <http://www.pisa.gc.ca/pisa/81-590-xpe.pdf#search=%22pisa%202000%20canada%22>

Bussiere, P. Cartwright, F. & Knighton, T (2004): *Measuring up: Canadian Results of the OECD PISA Study. The Performance of Canada's Youth in Mathematics, Reading, Science and Problem Solving. 2003 First Findings for Canadians Aged 15*. Ottawa, ON: Statistics Canada. Se: <http://www.pisa.gc.ca/81-590-xie2004001.pdf>

Dale, Erling Lars og Wærness, Jarl Inge (2006): *Vurdering og læring i en elevaktiv skole*. Oslo: Cappelen.

Dale, Erling Lars og Wærness, Jarl Inge (2003): *Sammenhengen mellom eksamen og læring. Læringsmiljø og læringsresultater. Avsluttende rapport*. Oslo: Læringscenteret.

Dale, Erling Lars (1999): *Utdanning med Pedagogisk Profesjonalitet*. Oslo: Ad Notam Gyldendal.

Det kongelige kirke-, utdannings- og forskningsdepartement (1996): *Læreplanverket for den 10-årige grunnskolen*. Oslo: Nasjonalt læremiddelsenter.

Engelsen, Britt Ulstrup (2003): *Ideer som formet vår skole? Læreplanen som idébærer – Et historisk perspektiv*. Oslo: Gyldendal Akademisk

Engelsen, Britt Ulstrup (2006): *Kan læring planlegges? Arbeid med læreplaner – hva, hvordan, hvorfor*. Oslo: Gyldendal Akademisk

Gregory, Katheleen, Cameron, Caren & Davies, Anne (2002): *Setting and Using Criteria. For use in Middle and Secondary School Classrooms*. British Colombia: Connections Publishing

Gundem, Bjørg B (1998) *Skolens oppgave og innhold. En studiebok i didaktikk*. Oslo: Universitetsforlaget

Gundem, Bjørg B og Hopman, Stefan (2002): *Didaktikk and / or Curriculum. An international Dialouge*. New York: Peter Lang Publishing

Haugstveit, Tove Brit (2005): Vurdering som profesjonsfaglig kompetanse – læreres refleksjoner over egen vurderingspraksis på 5, 6 og 7 trinn. *Samspillet mellom egenvurdering, selvoppfatning og prestasjoner er et felt det antakelig er nyttig for lærerne å utvikle videre innsikt*. I *Pedagogisk tidskrift* 6/ 2005 – 423.

Haugstveit, Tove Brit, Sjølie, Gunvor & Øygarden, Bjarne (2006): *Vurdering som profesjonsfaglig kompetanse. Rapport fra et KUPP prosjekt*. Høgskolen i Hedemark. Rapport nr 5.

Harrison, T. W & Kachur, J. L (1999): *Contested Classrooms. Education, Globalization and Democracy in Alberta. Canada*: The University of Alberta Press and Parkland institute.

Harvey, David (2005). *A brief history of neoliberalism*. New York: Oxford University Press.

Henmoe, Ola (2006): *Foreslår å bedømme elevene. Nytt system skal testes fra 2. Klasse*. I *Dagsavisen* 05.05.06

Hoel, Torlaug Løkensgard (2005): KUPP – programmet og behov for profesjonsforskning I *Pedagogisk tidskrift* 6/ 2005 – 423.I

Howell, Signe og Melhuus, Marit (2001): *Fjern og Nær – sosialantropologiske perspektiver på verdens samfunn og kulturer*. Oslo: Gyldendal Norsk Forlag

Kjærnsli, Marit og Svein Lie, Rolf Vegar Olsen, Astrid Roe og Are Turmoe (2004): *Rett spor eller ville veier? Norske elevers prestasjoner i matematikk, naturfag og lesing i PISA 2003*. Oslo. Universitetsforlaget.

Klette, Kirsti (2003): *Lærenes klasseromsarbeid; Interaksjons- og arbeidsformer i norske klasserom etter Reform 97*. I: Klette, Kirsti (red.): *Klasserommets praksisformer etter Reform 97*. Oslo: Pedagogisk forskningsinstitutt, Universitetet i Oslo.

Klein, N. (2001) *No Logo*, London: Flamingo

KUF (1998): *Elevvurdering. Rettledning til Læreplanverket for den 10 årige grunnskolen (L97)*. Oslo: Kirke-, utdannings- og forskningsdepartementet

KUF (1997): *Læreplanverket for den 10 årige grunnskolen (L97)*. Oslo: Kirke-, utdannings- og forskningsdepartementet

Kvale (2001): *Det kvalitative forskningsintervju*. Oslo: Gyldendal Norsk Forlag.

Lidén, Hilde (2001): Underforstått likhet - Skolens håndtering av forskjeller i et flerkulturelt samfunn. I: Vike, Halvard (red) 2001. *Likhetens paradokser - Antropologiske undersøkelser i det moderne Norge*. Oslo: Universitetsforlaget

Le Métais, Johanna (1999):” Chapter 6: Values and Aims in Curriculum and Assessment Frameworks: A 16 – Nation Review”, I (Red.) Moon, Bob & Murphy Patricia. *Curriculum in context*. Open University Press 1999 ss 93 – 113

Lund, T. (2002): Innledning. I: Lund, T. (red): *Innføring i forskningsmetodologi*. Oslo: Unipub forlag.

Lysne, A. (1999): *Karakterer og kompetanse. Stridstema i norsk skolehistorie*. Bind 1: Utviklingen fram til 1970-årene. AVA Forlag.

Mazurek Kas (1999): *Passing Fancies: Educational Changes in Alberta*. I Harrison, T. W & Kachur, J. L (red). *Contested Classrooms*. Education, Globalization and Democracy in Alberta. Canada: The University of Alberta Press and Parkland institute.

Nygård, Roald (2003): *Aktør eller Brikke? – om menneskers selvforståelse*. Oslo: Gyldendal norsk forlag

OECD (2005a): *Education at Glance*. OECD indicators 2005. OECD Publishing

OECD (2005b): *Formative Assessment. Improving learning in secondary classrooms*.

Roberts, Sylvia M & Pruitt, Eunice Z (2003): *Schools as Professional Learning Communities*. U.S: Corwin Press.

Pervin, Lawrence A & John, Oliver P John. (2001): *Personality. Theory and research*. USA: John Wiley & Sons.

Rubin, Herbert J & Rubin Irene S. (1995): *Qualitative Interviewing The art of Hearing Data*. London: Sage Publications

Scriven (1967): Chapter 5: The methodology of Evaluation. I (Red) Peter A. Taylor & Doris M. Cowley (1973) *Reading in Curriculum Evaluation*. Dubuque, Iowa: WM. C. Brown Company Publishers

Senge, Peter M (2004): *Den femte disiplin. Kunsten å utvikle den lærende organisasjon*. Oversatt av Arild Lillebø. Oslo: Pensumtjeneste A/S

Skagen (2000): Norsk veiledningstradisjon i lærerutdanning – analyse og kritikk, i (red). *Kunnskap og handling i pedagogisk veiledning*. Fagbokforlaget 2000. ss 10-34

St. meld. Nr 30 (2003-2004): *"Kultur for læring"*. Det kongelige kirke-, utdannings- og forskningsdepartement. Oslo: Nasjonalt læremiddelsenter.

Søgnen, Astrid (2002): *Førsteklasses fra Første klasse. Forslag til rammeverk for et nasjonalt kvalitetsvurderingssystem av norsk grunnopplæring*. NOU 2002:10. Oslo: Statens forvaltningstjeneste.

Telhaug, Alfred Oftedal (1994): *Utdanningspolitikken og enhetsskolen: studier i 1990-årenes utdanningspolitikk*. Oslo: Didakta

Telhaug, Alfred Oftedal (1997): *Utdanningsreformene: Oversikt og analyse*. Oslo: Didakta

UFD (2005): *Lærere elevene mer på lærende skoler? En snarvei til Kompetanseberetningen for Norge 2005*. 2. Utg. Oslo: Utdannings- og forskningsdepartementet

UFD (2002): *I Første rekke* (NOU: 16). Oslo: Utdannings og forskningsdepartementet

Utdanningsdirektoratet (2006): Elevvurdering i Kunnskapsløftet. Se: http://www.utdanningsdirektoratet.no/upload/Brev/Elevvurdering_i_Kunnskapsloftet.pdf

Utdanningsetaten i Oslo Kommune (2006): *Vurdering, Åpenhet og oppfølging – et helhetlig system for evaluering og vurdering i Utdanningsetaten*. Oslo: Utdanningsetaten i Oslo Kommune.

Vedung, Evert (1998): *Utvärdering i politik och forvaltning*. 2 utg. Lund: Studentlitteratur.

Wellington, Jerry (2000): *Educational research. Conemporary Issues and Practical approaches*. London: Continium

Øygarden, Bjarne (2006): Vurdering i inkluderende klasserom. I Haugstveit, Tove Brit, Sjølie, Gunvor & Øygarden, Bjarne (red). *Vurdering som profesjonsfaglig kompetanse. Rapport fra et KUPP prosjekt*. Høgskolen i Hedemark. Rapport nr 5.

Digitale referanser:

<http://forbruker.no/jobbogstudier/studier/article929687.ece>

www.education.gov.ab.ca

http://education.gov.ab.ca/k_12/curriculum/bySubject/

<http://www.epsb.ca/index.shtml>

www.infed.org/thinkers/senge/htm

http://odin.dep.no/kd/norsk/dok/andre_dok/nou/045001-020002/hov002-bn.html

<http://www.pisa.gc.ca/pisa/81-590-xpe.pdf#search=%22pisa%202000%20canada%22>

<http://www.pisa.gc.ca/81-590-xie2004001.pdf>

<http://www.united-online.ca/phpBB2/viewtopic.php?t=626&sid=81d27a5ee6ad46b33b2274918a7ee7b0>

www.utdanningsdirektoratet.no/templates/udir/TM_Artikkel.aspx?id=1400

http://www.utdanningsdirektoratet.no/upload/Rapporter/elevinspektorene/Som_elevene_ser_det_revidert_Elevinspektorene.pdf

<http://www.utdanningsetaten.oslo.kommune.no/>

http://www.utdanningsforbundet.no/UdfTemplates/Page_____30737.aspx

<http://en.wikipedia.org/wiki/Assessment>

<http://en.wikipedia.org/wiki/Evaluation>